

WHAT IS CLAIMED IS:

1. A method of determining primary sites from biomedical images, comprising:
 - identifying, by a computing system, a first biomedical image of a first sample from one of a primary site or a secondary site associated with a condition in a first subject;
 - applying, by the computing system, the first biomedical image to a site prediction model comprising a plurality of weights to determine the primary site for the condition, the site prediction model trained using a training dataset having a plurality of examples, each example comprising:
 - a second biomedical image of a second sample from one of the primary site or the secondary site of the condition in a second subject from which the second sample is obtained,
 - a first label identifying one of the primary site or the secondary site for the second sample from which the second biomedical image is obtained, and
 - a second label identifying the primary site for the condition in the second subject; and
 - storing, by the computing system, an association between the first biomedical image and the primary site determined using the site prediction model.
2. The method of claim 1, further comprising providing, by the computing system, the association between the first biomedical image and the primary site.
3. The method of claim 1, wherein applying further comprises applying the first biomedical image to the site prediction model to determine a plurality of candidate primary sites for the condition in the first subject.
4. The method of claim 1, wherein applying further comprises applying the first biomedical image to the site prediction model to determine a confidence score for the primary site for the condition.

5. The method of claim 1, wherein applying further comprises applying the first biomedical image to the site prediction model to determine a ranking for the plurality of candidate primary sites.
6. The method of claim 1, wherein identifying the first biomedical image further comprises obtaining the first biomedical image of the first sample via a histological image preparer.
7. The method of claim 1, wherein the plurality of weights in the site prediction model is arranged into (i) a plurality of convolution blocks to generate a plurality of feature maps from the biomedical image and (ii) an activation layer to determine the primary site for the condition based on the plurality of feature maps.
8. A method of training models to determine primary sites from biomedical images, comprising:
 - identifying, by a computing system, a training dataset having a plurality of examples, each example of the plurality of examples comprising:
 - a biomedical image of a sample from one of the primary site or the secondary site of the condition in a second subject from which the sample is obtained,
 - a first label identifying one of the primary site or the secondary site for the second sample from which the biomedical image is obtained, and
 - a second label identifying the primary site for the condition in the second subject;
 - applying, by the computing system, the biomedical image in each of the plurality of examples of the training dataset to a site prediction model comprising a plurality of weights to determine a site for the condition of the sample;
 - comparing, by the computing system, for each example of the plurality of examples in the training dataset, the primary site identified in the label of the example and the site determined by the site prediction model;
 - updating, by the computing system, at least one of the plurality of weights in the site prediction model based on the comparison between the first site identified in the label of each example and the second site determined by the site prediction model; and

storing, by the computing system, in one or more data structures, the plurality of weights in the site prediction model.

9. The method of claim 8, further comprising applying, by the computing system, an acquired biomedical image of a second sample to the site prediction model to determine a second site for the second sample.

10. The method of claim 8, further comprising reapplying, by the computing system, the biomedical image of at least one example of the plurality of examples to the site prediction model, responsive to determining that a loss metric for the at least one example exceeds a threshold.

11. The method of claim 8, wherein updating further comprises updating at least one of the plurality of weights in the site prediction model using a classification loss determined based on the comparison.

12. The method of claim 8, wherein applying further comprises applying the biomedical image in each of the plurality of examples of the training dataset to determine a plurality of candidate primary sites for the condition in the first subject.

13. The method of claim 8, wherein applying further comprises applying the biomedical image in each of the plurality of examples of the training dataset to determine a confidence score for the primary site for the condition.

14. The method of claim 8, wherein the plurality of weights in the site prediction model is arranged into (i) a plurality of convolution blocks to generate a plurality of feature maps from the biomedical image and (ii) an activation layer to determine the primary site for the condition based on the plurality of feature maps.

15. A system for determining primary sites from biomedical images of metastatic sites, comprising:

a computing system having one or more processors coupled with memory, configured to:

- identify a first biomedical image of a first sample from one of a primary site or a secondary site associated with a condition in a first subject;

- apply the biomedical image to a site prediction model comprising a plurality of weights to determine the primary site for the condition, the site prediction model trained using a training dataset having a plurality of examples, each example comprising:

- a second biomedical image of a second sample from one of the primary site or the secondary site of the condition in a second subject from which the second sample is obtained,

- a first label identifying one of the primary site or the secondary site for the second sample from which the second biomedical image is obtained, and

- a second label identifying the primary site for the condition in the second subject; and

- store an association between the first biomedical image and the primary site determined using the site prediction model.

16. The system of claim 15, wherein the computing system is further configured to provide the association between the first biomedical image and the primary site.

17. The system of claim 15, wherein the computing system is further configured to apply the first biomedical image to the site prediction model to determine a plurality of candidate primary sites for the condition in the first subject.

18. The system of claim 15, wherein the computing system is further configured to apply the first biomedical image to the site prediction model to determine a confidence score for the primary site for the condition.

19. The system of claim 15, wherein the computing system is further configured to apply the first biomedical image to the site prediction model to determine a ranking for the plurality of candidate primary sites.

20. The system of claim 15, wherein the plurality of weights in the site prediction model is arranged into (i) a plurality of convolution blocks to generate a plurality of feature maps from the biomedical image and (ii) an activation layer to determine the primary site for the condition based on the plurality of feature maps.