medium. Data storage device **1312**, and memory **1310**, may each include high-speed random access memory, such as dynamic random access memory (DRAM), static random access memory (SRAM), double data rate synchronous dynamic random access memory (DDR RAM), or other random access solid state memory devices, and may include non-volatile memory, such as one or more magnetic disk storage devices such as internal hard disks and removable disks, magneto-optical disk storage devices, optical disk storage devices, flash memory devices, semiconductor memory devices, such as erasable programmable read-only memory (EPROM), electrically erasable programmable read-only memory (EEPROM), compact disc read-only memory (CD-ROM), digital versatile disc read-only memory (DVD-ROM) disks, or other non-volatile solid state storage devices.

[0163] Input/output devices **1308** may include peripherals, such as a printer, scanner, display screen, etc. For example, input/output devices **1308** may include a display device such as a cathode ray tube (CRT) or liquid crystal display (LCD) monitor for displaying information to the user, a keyboard, and a pointing device such as a mouse or a trackball by which the user can provide input to computer **1302**.

[0164] An image acquisition device **1314** can be connected to the computer **1302** to input image data (e.g., medical images) to the computer **1302**. It is possible to implement the image acquisition device **1314** and the computer **1302** as one device. It is also possible that the image acquisition device **1314** and the computer **1302** communicate wirelessly through a network. In a possible embodiment, the computer **1302** can be located remotely with respect to the image acquisition device **1314**.

[0165] Any or all of the systems and apparatus discussed herein may be implemented using one or more computers such as computer **1302**.

[0166] One skilled in the art will recognize that an implementation of an actual computer or computer system may have other structures and may contain other components as well, and that FIG. **13** is a high level representation of some of the components of such a computer for illustrative purposes.

[0167] The foregoing Detailed Description is to be understood as being in every respect illustrative and exemplary, but not restrictive, and the scope of the invention disclosed herein is not to be determined from the Detailed Description, but rather from the claims as interpreted according to the full breadth permitted by the patent laws. It is to be understood that the embodiments shown and described herein are only illustrative of the principles of the present invention and that various modifications may be implemented by those skilled in the art without departing from the scope and spirit of the invention. Those skilled in the art could implement various other feature combinations without departing from the scope and spirit of the invention.

1. A method for classifying a lesion in a medical image, comprising:

receiving an input medical image depicting a lesion;

localizing the lesion in the input medical image using a trained localization network to generate a localization map;

classifying the lesion based on the input medical image and the localization map using a trained classification network; and

outputting the classification of the lesion,

wherein the trained localization network and the trained classification network are jointly trained.

2. The method of claim **1**, further comprising:

jointly training the trained localization network and the trained classification network by separately training the localization network to determine weights of the localization network during a first training phase, and training the classification network based on the weights of the localization network during a second training phase.

3. The method of claim **2**, wherein separately training the localization network comprises:

receiving a multi-site dataset associated with different clinical sites and a deployment dataset associated with a deployment clinical site;

training a deep learning model based on the multi-site dataset; and

optimizing the trained deep learning model based on the deployment dataset to provide the trained localization network.

4. The method of claim **1**, wherein:

the input medical image is a multi-parametric magnetic resonance imaging (mpMRI) image comprising a plurality of images;

localizing the lesion in the input medical image using a trained localization network to generate a localization map comprises localizing the lesion in each of the plurality of images using the trained localization network to generate a localization map for each of the plurality of images; and

classifying the lesion based on the input medical image and the localization map using a trained classification network comprises:

combining the localization maps for the plurality of images, and

classifying the lesion based on the plurality of images and the combined localization maps using the trained classification network.

5. The method of claim **4**, further comprising:

preprocessing the plurality of images of the mpMRI image to address variances between the plurality of images.

6. The method of claim **5**, wherein preprocessing the plurality of images of the mpMRI image to address variances between the plurality of images comprises:

removing geometric variability in the plurality of images of the mpMRI image.

7. The method of claim **5**, wherein preprocessing the plurality of images of the mpMRI image to address variances between the plurality of images comprises:

normalizing intensity variability in the plurality of images of the mpMRI image.

8. The method of claim **1**, wherein the localization map is associated with a score of the lesion.

9. The method of claim **1**, further comprising:

generating a clinical relevance map and a label indicating whether the lesion is clinically significant based on the localization map and one or more of: a size of the lesion, an average intensity within the lesion, a variance of intensities within the lesion, radiomic features, and lexicon based features.

10. The method of claim **1**, further comprising:

generating a patch from the input medical image using the localization map; and