What is claimed is:

1. A computer-implemented method for processing input activations for a neural network layer, the method, comprising:

receiving, by a computing device, a plurality of input activations, the input activations being provided, at least in part, from a source external to the computing device;

determining, by a controller of the computing device, whether each of the plurality of input activations has one of a zero value or a non-zero value;

storing, in a memory bank of the computing device, at least one of the input activations;

generating, by the controller, an index identifying only memory address locations in the memory bank that store non-zero input activation values; and

providing, by the controller and from the memory bank, at least one input activation onto a data bus that is accessible by one or more units of a computational array for performing computations based on the at least one input activation, wherein the at least one input activation is provided, at least in part, from a memory address location associated with the index.

2. The method of claim 1, wherein the index is created based on a bitmap comprising a plurality of bits and, wherein each bit of the bitmap indicates at least one of a non-zero input activation value or a zero input activation value.

3. The method of claim 1, further including, providing a first input activation that has a non-zero value to perform, by at least one unit, a computation using the non-zero value, and subsequently providing a second input activation that has a zero value, and preventing, in at least one unit, computation that would otherwise be performed using the zero value.

4. The method of claim 3, wherein preventing occurs in response to the controller determining that the input activation is provided from a memory address location that is not associated with the index.

5. The method of claim 4, further including, detecting, by the controller, that the input activation is provided from a memory address location that is not associated with the

index, and, in response to detecting, providing a control signal to at least one unit of the computational array to prevent a multiply operation associated with the zero input activation value.

6. The method of claim 1, wherein the method further comprises, mapping, by the controller and to a first unit, a first portion of a tensor computation that uses a first input activation and mapping, to a second unit that differs from the first unit, a second portion of the tensor computation that also uses the first input activation.

7. The method of claim 1, further comprising, sequentially providing a single input activation onto the data bus, the single input activation being accessed and selected from memory address locations that are associated with the index.

8. The method of claim 1, wherein providing further comprises, not providing input activations that have a zero value.

9. One or more machine-readable storage devices storing instructions for processing input activations for a neural network layer, the instructions being executable by one or more processing devices to perform operations comprising:

receiving, by a computing device, a plurality of input activations, the input activations being provided, at least in part, from a source external to the computing device;

determining, by a controller of the computing device, whether each of the plurality of input activations has one of a zero value or a non-zero value;

storing, in a memory bank of the computing device, at least one of the input activations;

generating, by the controller, an index identifying only memory address locations in the memory bank that store non-zero input activation values; and

providing, by the controller and from the memory bank, at least one input activation onto a data bus that is accessible by one or more units of a computational array for performing computations based on the at least one input activation, wherein the at least one input activation is provided, at least in part, from a memory address location associated with the index.

10. The machine-readable storage devices of claim 9, wherein the index is created based on a bitmap comprising a plurality of bits and, wherein each bit of the bitmap indicates at least one of a non-zero input activation value or a zero input activation value.

11. The machine-readable storage devices of claim 9, further including, providing a first input activation that has a non-zero value to perform, by at least one unit, a computation using the non-zero value, and subsequently providing a second input activation that has a zero value, and preventing, in at least one unit, computation that would otherwise be performed using the zero value.

12. The machine-readable storage devices of claim 11, wherein preventing occurs in response to the controller determining that the input activation is provided from a memory address location that is not associated with the index.

13. The machine-readable storage devices of claim 12, further including, detecting, by the controller, that the input activation is provided from a memory address location that is not associated with the index, and, in response to detecting, providing a control signal to at least one unit of the computational array to prevent a multiply operation associated with the zero input activation value.

14. The machine-readable storage devices of claim 9, wherein the operations further comprise, mapping, by the controller and to a first unit, a first portion of a tensor computation that uses a first input activation and mapping, to a second unit that differs from the first unit, a second portion of the tensor computation that also uses the first input activation.

15. An electronic system for processing input activations for a neural network layer, the system comprising:

a controller disposed in a computing device, the controller including one or more processing devices; and

one or more machine-readable storage devices for storing instructions that are executable by the one or more processing devices to perform operations comprising:

receiving, by the computing device, a plurality of input activations, the input activations being provided, at least in part, from a source external to the computing device;

determining, by the controller, whether each of the plurality of input activations has one of a zero value or a non-zero value;

storing, in a memory bank of the computing device, at least one of the input activations;

generating, by the controller, an index identifying only memory address locations in the memory bank that store non-zero input activation values; and

providing, by the controller and from the memory bank, at least one input activation onto a data bus that is accessible by one or more units of a computational array for performing computations based on the at least one input activation, wherein the at least one input activation is provided, at least in part, from a memory address location associated with the index.

16. The electronic system of claim 15, wherein the index is created based on a bitmap comprising a plurality of bits and, wherein each bit of the bitmap indicates at least one of a non-zero input activation value or a zero input activation value.

17. The electronic system of claim 15, further including, providing a first input activation that has a non-zero value to perform, by at least one unit, a computation using the non-zero value, and subsequently providing a second input activation that has a zero value, and preventing, in at least one unit, computation that would otherwise be performed using the zero value.

18. The electronic system of claim 17, wherein preventing occurs in response to the controller determining that the input activation is provided from a memory address location that is not associated with the index.

19. The electronic system of claim 17, further including, detecting, by the controller, that the input activation is provided from a memory address location that is not associated with the index, and, in response to detecting, providing a control signal to at least one unit of the computational array to prevent a multiply operation associated with the zero input activation value.

20. The electronic system of claim 15, wherein the operations further comprise, mapping, by the controller and to a first unit, a first portion of a tensor computation that uses a first input activation and mapping, to a second unit that differs from the first unit, a second portion of the tensor computation that also uses the first input activation.