LeNet architecture

- **Input Image (32x32x1):** The architecture starts with the input image, which is typically a grayscale image of size 32x32 pixels.
- **Convolutional Layer (6 filters, 5x5):** The first convolutional layer applies six 5x5 filters to the input image. Each filter extracts different features from the input image, creating six feature maps.
- **Pooling Layer (Average Pooling):** After each convolutional layer, a pooling layer is applied to reduce the spatial dimensions of the feature maps. In LeNet, average pooling with a kernel size of 2x2 is commonly used.
- **Convolutional Layer (16 filters, 5x5):** Another convolutional layer follows the pooling layer, applying 16 filters of size 5x5 to the feature maps generated from the previous layer.
- **Pooling Layer (Average Pooling):** Similar to before, average pooling is applied to reduce the spatial dimensions of the feature maps.

LeNet architecture

- **Flatten:** After the convolutional and pooling layers, the feature maps are flattened into a one-dimensional vector. This prepares the data for feeding into the fully connected layers.

- **Fully Connected Layer (120 neurons):** The flattened features are then passed through a fully connected layer with 120 neurons. Each neuron in this layer is connected to every element in the flattened feature vector.

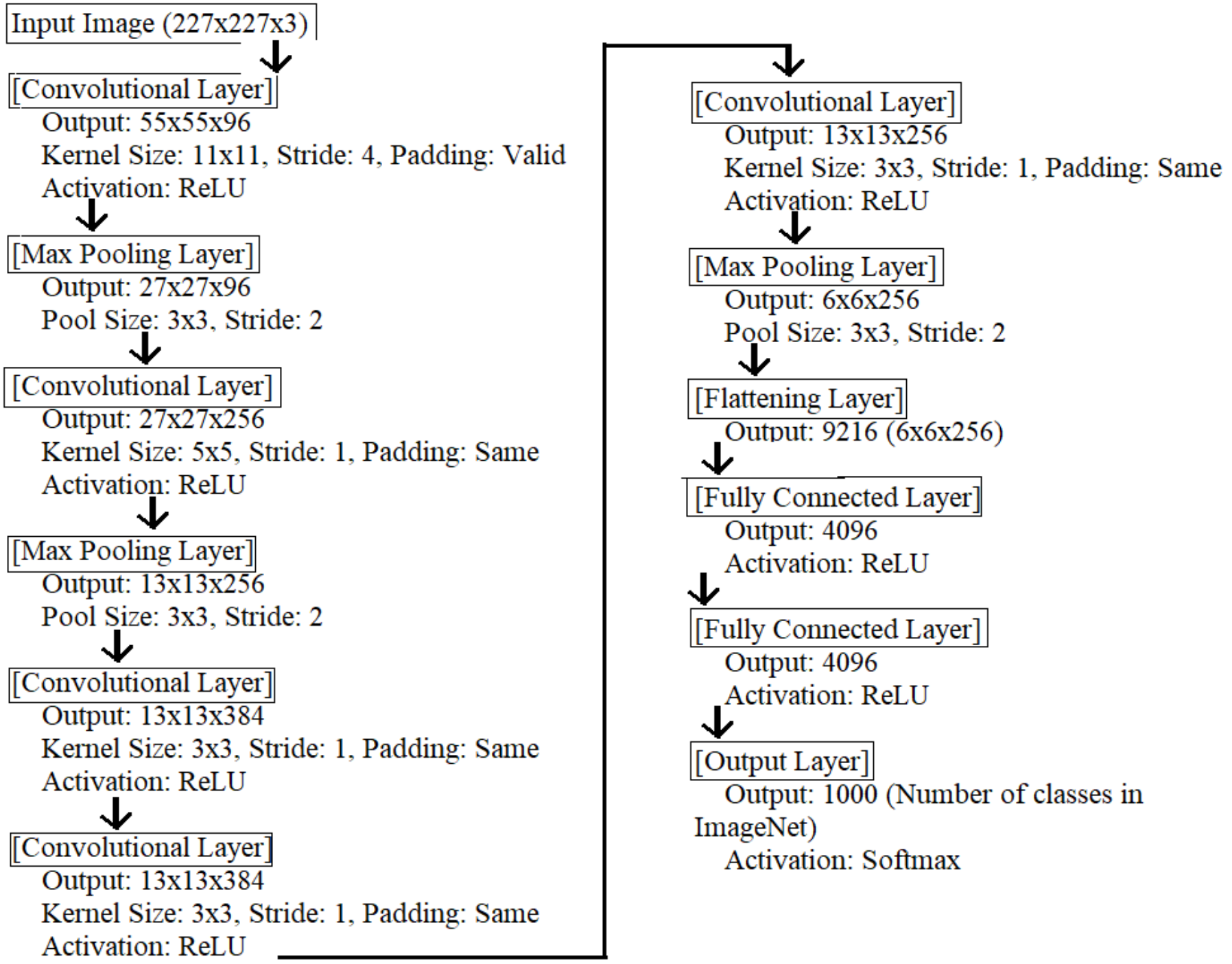
- **Fully Connected Layer (84 neurons):** Another fully connected layer with 84 neurons follows the previous layer.

- **Output Layer (10 neurons, Softmax):** The final layer of the LeNet architecture is the output layer, consisting of 10 neurons corresponding to the possible classes in the classification task. The softmax activation function is commonly used here to convert the raw scores into class probabilities.

- THE **OUTPUT LAYER** IS ALSO A FULLY CONNECTED LAYER, WITH A TOTAL OF 10 NODES, WHICH RESPECTIVELY REPRESENT THE NUMBERS 0 TO 9, AND IF THE VALUE OF NODE I IS 0, THE RESULT OF NETWORK RECOGNITION IS THE NUMBER I. A RADIAL BASIS FUNCTION (RBF) NETWORK CONNECTION IS USED.
- ASSUMING X IS THE INPUT OF THE PREVIOUS LAYER AND Y IS THE OUTPUT OF THE RBF, THE CALCULATION OF THE RBF OUTPUT IS:

$$y_i = \sum_j (x_j - w_{ij})^2$$

AlexNet Architecture



AlexNet Architecture

- AlexNet is a pioneering convolutional neural network (CNN) architecture that played a significant role in advancing the field of deep learning, particularly in computer vision tasks.
- It was developed by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton and won the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2012.

- **Input Layer:** Accepts input images of size 227x227x3 (RGB images).

- **Convolutional Layers:** AlexNet consists of five convolutional layers. Each convolutional layer is followed by a ReLU activation function. These layers learn hierarchical features from the input images.

- **Max Pooling Layers:** Four max-pooling layers are interspersed between the convolutional layers. They downsample the feature maps to reduce spatial dimensions while retaining important information.

AlexNet Architecture

- **Fully Connected Layers:** Following the convolutional and pooling layers are three fully connected layers. These layers integrate the learned features and perform classification based on the extracted features.

- **Output Layer:** The final layer is a softmax activation layer that produces class probabilities. In the original AlexNet, it outputs probabilities for 1000 classes from the ImageNet dataset.

- **Dropout:** Dropout regularization is applied before the fully connected layers to prevent overfitting by randomly dropping neurons during training.

- AlexNet demonstrated the effectiveness of deep learning in image classification tasks and paved the way for subsequent advancements in CNN architectures.