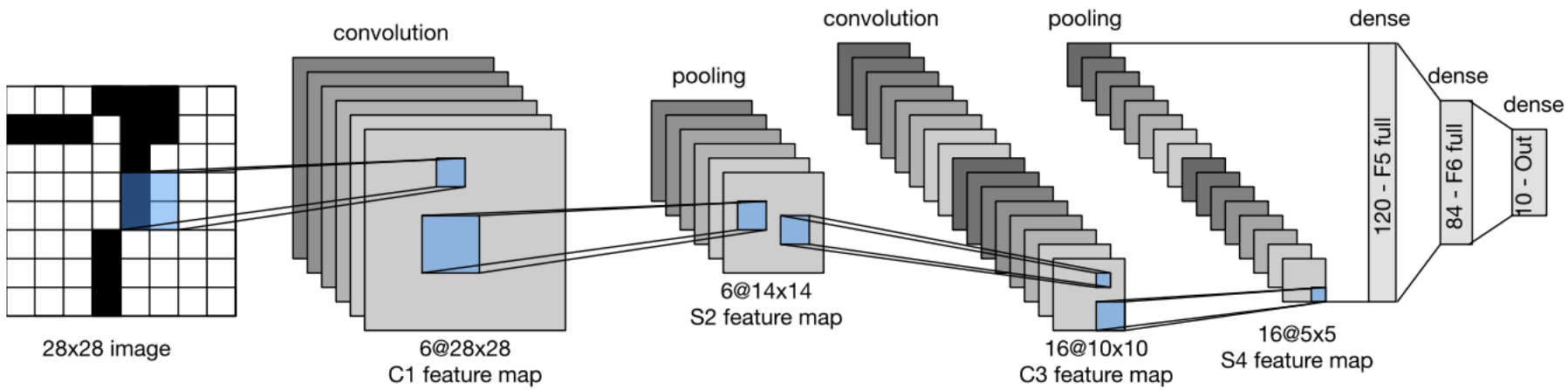
Convolutional Neural Networks – LeNet

- **LeNet** is the first published CNNs
- The model was introduced by Yann LeCun, then a researcher at AT&T Bell Labs, for recognizing handwritten digits in images
- In 1989, LeCun published the first study to successfully train CNNs via backpropagation.
- At the time LeNet achieved outstanding results matching the performance of support vector machines, then a dominant approach in supervised learning.
- LeNet was eventually adapted to recognize digits for processing deposits in ATM machines.

LeNet

- At a high level, LeNet (LeNet-5) consists of 2 parts:
    1. a convolutional encoder consisting of two convolutional layers; and
    2. a dense block consisting of three fully-connected layers;



LeNet – Convolutional Encoder

- Each convolutional *block*:
  - A convolutional layer.
  - A sigmoid activation function (ReLUs were discovered recently).
  - A subsequent average pooling operation (max pooling was discovered later).
- Each convolutional layer uses a 5×5 kernel.
- The first convolutional layer has 6 output channels, while the second has 16.
- Each 2×2 pooling operation (stride 2) reduces dimensionality by a factor of 4 via spatial downsampling.
- The convolutional block emits an output with shape given by (batch size, number of channel, height, width).

LeNet – Dense Block

- In order to pass output from the convolutional block to the dense block, we must flatten each example in the minibatch.
- In other words, we take the four-dimensional input and transform it into the two-dimensional input expected by fully-connected layers:
    - the two-dimensional representation that we desire has uses the first dimension to index examples in the minibatch
    - the second to give the flat vector representation of each example.
- LeNet's dense block has three fully-connected layers, with 120, 84, and 10 outputs, respectively.
    - Because we are still performing classification, the 10-dimensional output layer corresponds to the number of possible output classes.

Compressed LeNet Representation