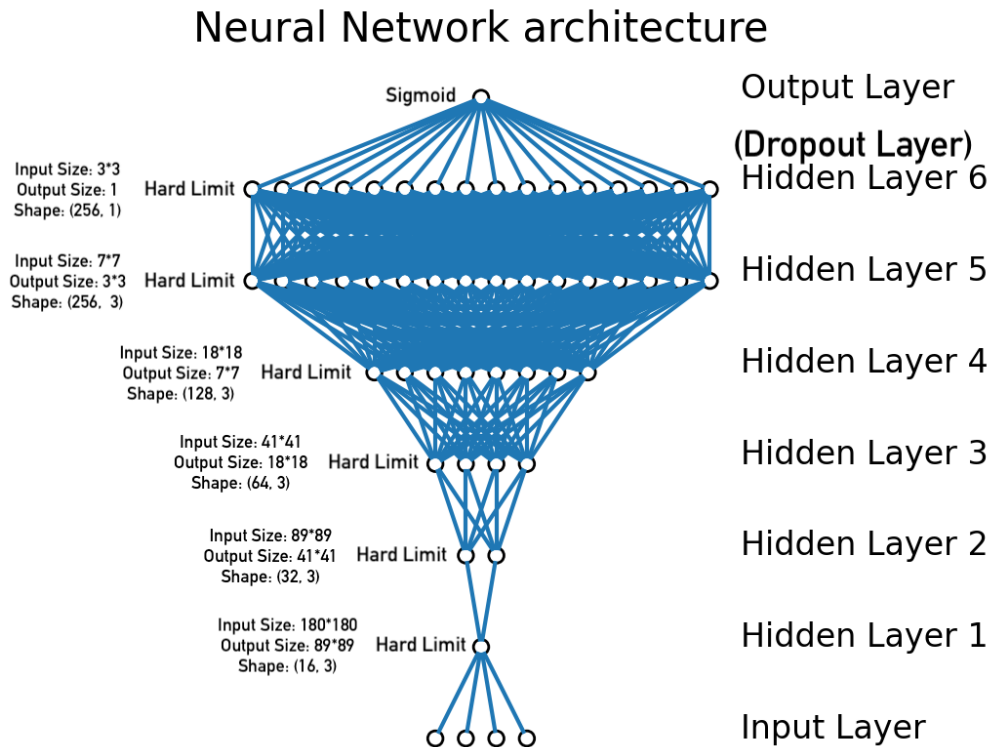


- Design, input sizes, output sizes, shapes, activation functions, and dropout layer:



- Reason of usage of the activation functions and dropout layer:

- Activation functions:

- 1) We have used a hard-limit function ( $\text{input} \leq 0 \rightarrow 0$ ,  $\text{input} > 0 \rightarrow \text{input}$ ) in all of the hidden layers because it is very computationally efficient to compute, where it just needs to output one of two discrete numbers (0 or input), and our model has tons of neurons in the hidden layers so we needed our activation function to be as simple and computationally efficient as possible.
- 2) We have used a sigmoid function ( $\rightarrow 1/(1+e^{(-\text{input})})$ ) in the output layer in order to output a probability prediction (0 to 1) for our image instead of a discrete one (0 or 1) in order to be more precise in our predictions due to the sensitivity of the given task.

- Dropout layer:

We have used a dropout layer directly before the output layer in order to drop 0.1 of our input units when predicting new data, in order to prevent our model from overfitting on the training data.