Neural network learning has achieved remarkable success in a wide range of real-world domains, such as image recognition and natural language processing. However, in dealing with functional variables such as curves and surfaces/images, the interrelationship between discrete observations and the intrinsic smoothness of functions are not fully captured. To fill this gap, we propose the Functional Universal Approximation (FUA) theorem for functional data which takes the whole curve or surface as an observation unit and can fully use the intrinsic smoothness feature of the curves or surfaces. With the FUA theorem, we consider a general nonlinear function-on-function (FOF) regression model which does not make any specific assumption on the form of the regression between functional variables, and develop a novel method, the fully Functional Neural Network (FNN) with one smooth hidden layer, to predict the functional response from functional predictors. All the components in the fully FNN are smooth functions, and we can naturally control the complexity of the FNN by controlling the smoothness of its component functions with smoothness regularization.