

## Residual Networks (Part 2)

Dr. Alireza Aghamohammadi

## Motivation

- Attempt to create a model that accurately represents the difference between the input and the output.



# Residual Network

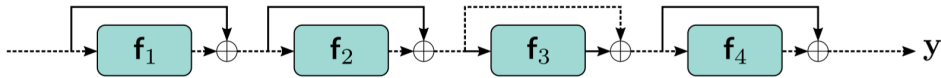
- A residual network is structured such that the input to each layer is added to its output. This is represented by the following equations<sup>1</sup>:

$$h_1 = x + f_1[x, \phi_1]$$

$$h_2 = h_1 + f_2[h_1, \phi_2]$$

$$h_3 = h_2 + f_3[h_2, \phi_3]$$

$$y = h_3 + f_4[h_3, \phi_4]$$



---

<sup>1</sup>Reference: "Understanding Deep Learning"

## Residual Block

- ▶ A residual block is a group of layers designed such that data flows through the layers and also bypasses them via skip connections.
- ▶ The symbol  $\oplus$  denotes an operation that either adds or concatenates elements to form a tensor.

