Quiz 8 Name: Student ID:

Q1. (3pts) Briefly explain the vanishing gradient problem and why Relu in AlexNet can help mitigate the problem

Q2. (3pts) Why GoogleNet (Inception) uses multiple variable size kernels in the same layer and how do they reduce the number of parameters from multiple kernels?

Q3. (4pt) Use the figure below to 1) briefly explain what residual connections do, 2) unravel the networks into shallow networks (show how X3 can be represented by X0 and functions in the networks)

Diagram

Description automatically generated