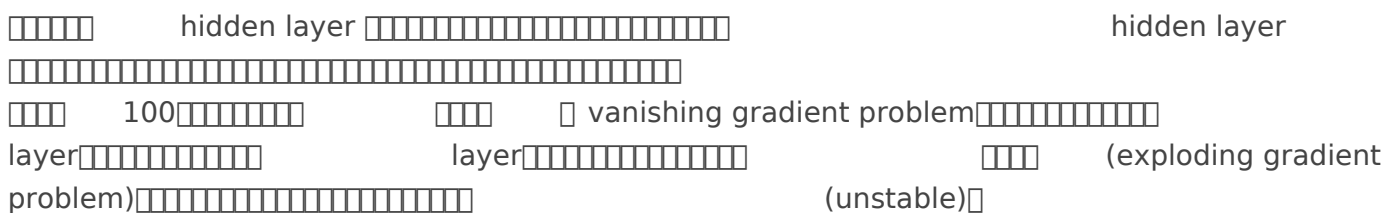

























Batch Normalization



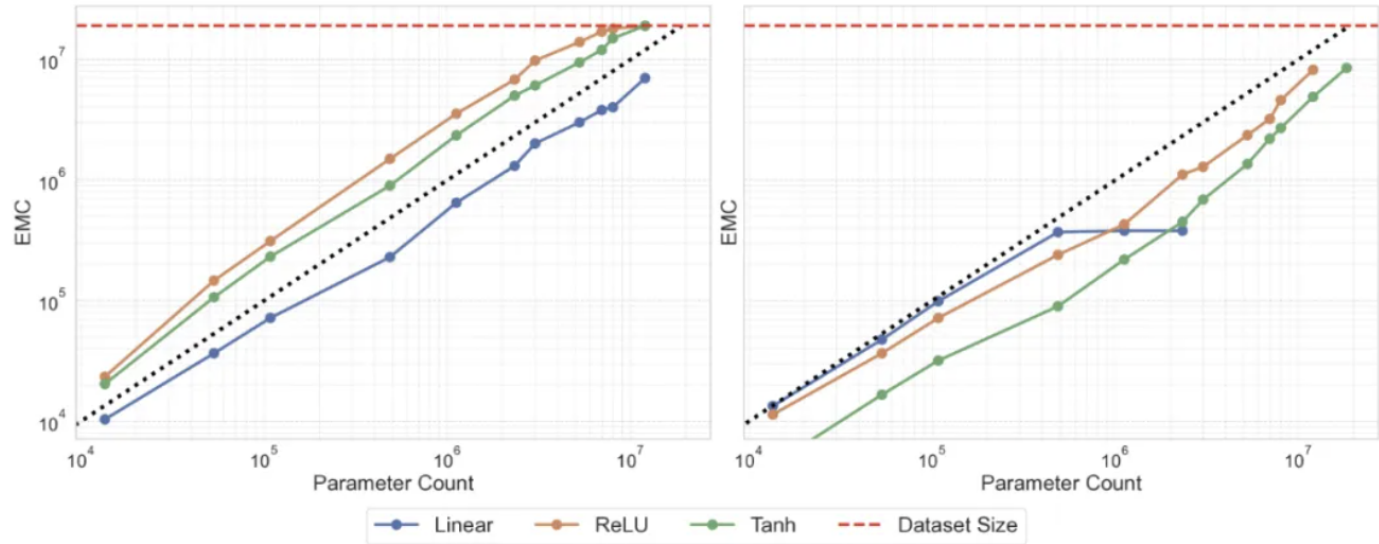
- 
-  **bias**  0  Batch Normalization  Standardization 
-   0  Batch Normalization  
-     
-   **sigmoid**  **tanh**  BN 
- **Batch Normalization**  **dropout** 



ReLU



ReLU



(a) Original Labels

(b) Random Labels

Figure 16: **ReLU networks exhibit higher flexibility.** EMC as a function of the number of parameters across different activation functions for original labels **(left)** and for random ones **(right)** on ImageNet-20MS.

公众号·新智元

• Sigmoid

-
-
-

sigmoid

image.png

image.png

• tanh

-

image.pngimage.png

• RELU

-
-
-
-

Revision #1

Created 11 January 2025 09:44:04 by Colin

Updated 11 January 2025 09:44:06 by Colin