Network Structures and ConvNet

Paper 1: A Battle of Network Structures: An Empirical Study of CNN, Transformer, and MLP

Paper 2: A ConvNet for the 2020s

1. Describe the SPACH framework adopted in paper 1. What does the author conclude about different network performance behavior at different scales with this framework?

2. Discuss briefly all four findings presented in paper 1.

3. Researchers when trying to design a standalone transformer network for vision ta sks face certain difficulties. State two such difficulties discussed in paper 1.

4. Describe the role of spatial mixing function and channel mixing function discussed in paper 1.

5. Paper 1 uses three distinct implementations of spatial mixing function (fs). Describe each of them.

6. As discussed in paper 1, MLP based models are prone to over-fitting. Describe two methods that are used for alleviating this problem.

7. Define translation invariance and translation equivariance concerning the CNNs. What helps a CNN achieve translation equivariance and translation invariance?

8. Describe in detail the series of design decisions applied on the baseline model in p aper 2. List and discuss all the design components changed under each of these five headers.

9. Briefly describe the following: (a) Local modeling (paper 1) (b) Hybrid models (paper 1) (c) Architectural difference between single-stage and multi-stage SPACH frameworks (paper 1) (d) Inductive bias (in ML models) (e) Isotropic model design