Are deep networks broken, and should we fix them?

Abstract:

We analyze three cases where Deep Neural Networks (DNNs) seem to work "sub-optimally". We find if these issues can-or should-be fixed.

Convnets were originally designed to be shift-invariant, but this does not hold because of aliasing. We show how to completely fix this issue using polynomial activations and achieve state-of-the-art performance under adversarial shifts-even for fractional shifts.

DNNs are known to exhibit "catastrophic forgetting" when trained on sequential tasks. We show this can happen even in a linear setting for regression and classification. However, we derive universal bounds which guarantee no catastrophic forgetting in certain cases, such as when tasks are repeated or randomly ordered.

It was recently observed that DNNs, optimized with gradient descent, operate at the "Edge of Stability", in which monotone convergence is not guaranteed. However, we prove a different potential function ("gradient flow solution sharpness") is monotonically decreasing in scalar networks, observe this holds empirically in DNNs, and discuss the implications.

References:

