Deep Feature Transmission Simulator v2 for Collaborative Intelligence

TensorFlow 2, Tensor Completion & Error Concealment

Hans Dhondea

Supervisor: Prof. Ivan Bajić

Multimedia Lab
School of Engineering Science
Simon Fraser University

October 2, 2020

Overview

Collaborative Intelligence

DFTS

Tensor Completion

Error Concealment

Future work

Introduction

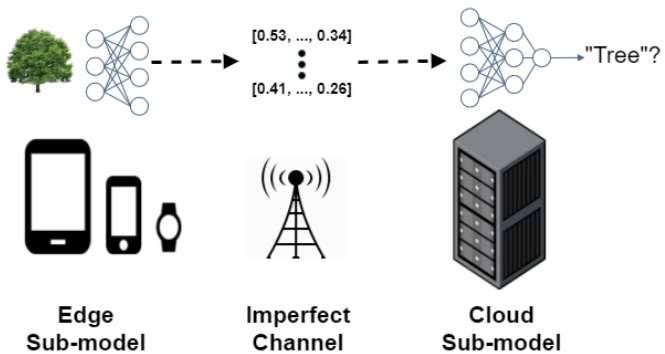


Figure 1: Blueprint for Collaborative Intelligence.[1] [1]

[1] Y. Kang, J. Hauswald, C. Gao, A. Rovinski, T. Mudge,J. Mars, and L. Tang, "Neurosurgeon: Collaborative intelligence between the cloud and mobile edge," SIGARCH Comput. Archit. News, vol. 45, p. 615–629, Apr. 2017.

Splitting a DNN

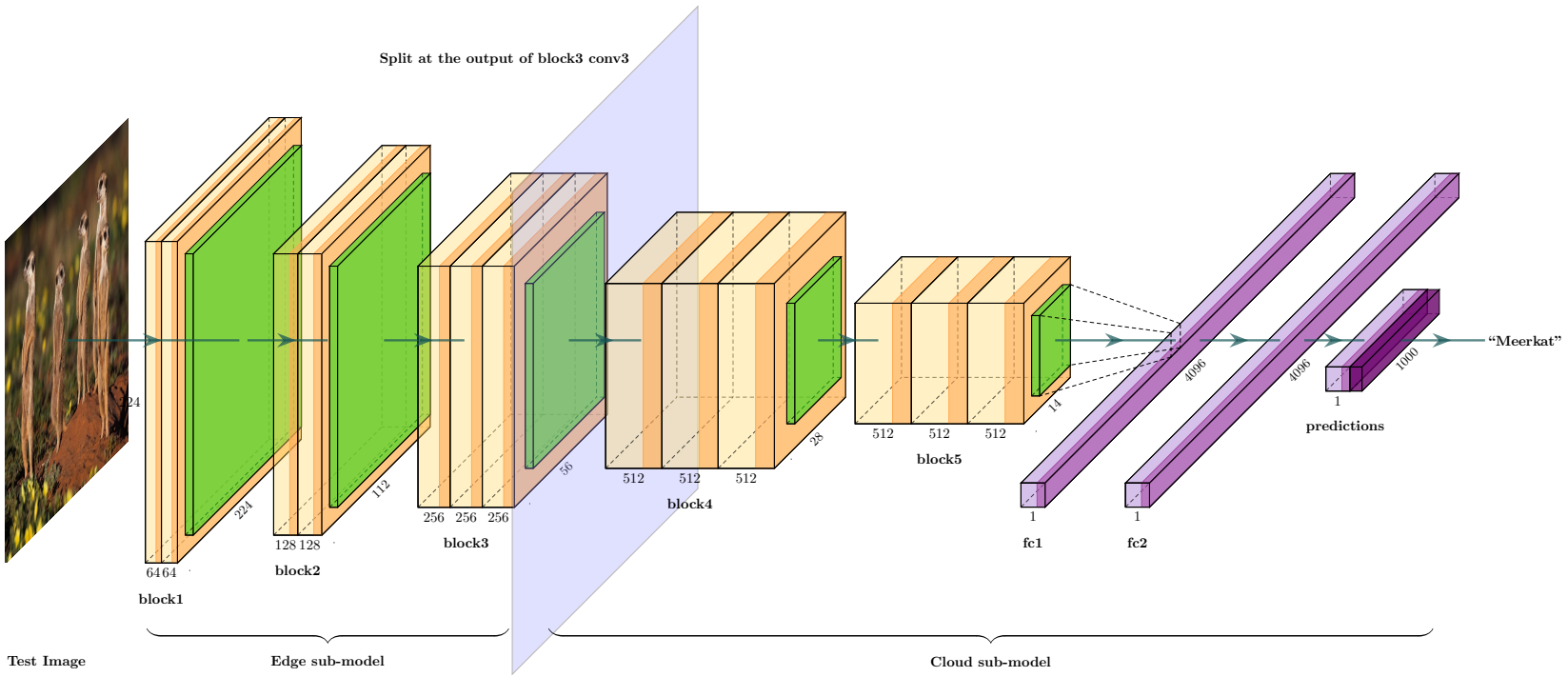


Figure 2: Splitting a deep model into two sub-models: the mobile or device sub-model and the remote or cloud sub-model.

Deep Feature Transmission Simulator (DFTS)

- DFTS was developed to simulate packet-based transmission of deep feature over unreliable communication channels [2] [2].
- DFTS simulations can include a Gilbert-Elliot channel, a random loss channel or a perfect channel (packets never incur any damage).
- n bit quantization can be applied to deep feature packets before transmission.
- To conceal errors in damaged tensors, DFTS can do linear interpolation and nearest neighbor interpolation.

[2] H. Unnibhavi, H. Choi, S. R. Alvar, and I. V. Bajić, "Dfts: Deep feature transmission simulator," 2018.

Updating DFTS

- DFTS was developed with TensorFlow 1.1.2 and Keras 2.2.2.
- Various changes in TensorFlow version 2 (discussed in the *Effective Tensorflow 2* guide[3] [3]) break the operation of DFTS.
- It is possible to run DFTS version 1 in TensorFlow 2 by disabling the v2 behavior.
- A number of higher-level API calls in DFTS were modified to become TF 2 compatible.

[3] `https://www.tensorflow.org/guide/effective_tf2`

Splitting a DNN into a mobile and a cloud client

Packetization

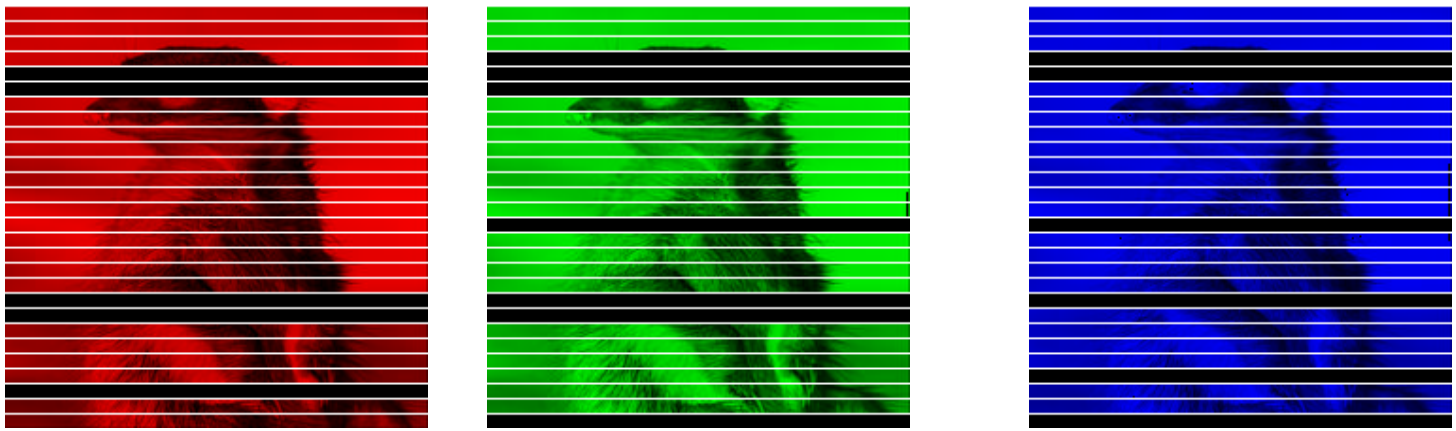
Channel model

- The channel model aims to. Three possibilities: no channel, Gilbert-Elliott channel & Random Loss channel.
- The Gilbert-Elliott model has been shown to fittingly capture real-time observed packet loss patterns over real-time services over the Internet [4]. [4].

[4] G. Hasslinger and O. Hohlfeld, "The gilbert-elliott model for packet loss in real time services on the internet," in 14th GI/ITG Conference - Measurement, Modelling and Evalutation of Computer and Communication Systems, pp. 1–15, 2008.

The Gilbert-Elliott channel model

- 8 rows per packet, so a $224 \times 224 \times 3$ tensor is packetized into 28 packets of 8×224 features per channel.
- Gilbert-Elliott channel of loss probability $P_B = 10\%$ and average burst length of $L_B = 4$ packets.



(a) Channel # 0. (b) Channel # 1. (c) Channel # 2

Figure 3: The RGB channels of an image packetized and transmitted through a Gilbert-Elliott channel.

Running DFTS simulations

- The *BrokenModel* class imports Tensorflow, which is best run on GPU.
- The *Channel* error concealment and tensor completion code are best run on CPU.

Tensor Completion overview & example

How many iterations are sufficient?

Error Concealment overview

Recommendations for future work

- Public repo for DFTSv2: `https://github.com/AshivDhondea/DFTS_TF2`.
- The object detection task has yet to be implemented.
- Run ALTeC and other tensor completion methods in a DFTS experiment.
- Run tensor completion and error concealment methods with speed-matching.

- Y. Kang, J. Hauswald, C. Gao, A. Rovinski, T. Mudge, J. Mars, and L. Tang, "Neurosurgeon: Collaborative intelligence between the cloud and mobile edge," *SIGARCH Comput. Archit. News*, vol. 45, p. 615–629, Apr. 2017.
- H. Unnibhavi, H. Choi, S. R. Alvar, and I. V. Bajić, "Dfts: Deep feature transmission simulator," 2018.
- "Effective tensorflow 2."
- G. Hasslinger and O. Hohlfeld, "The gilbert-elliott model for packet loss in real time services on the internet," in *14th GI/ITG Conference - Measurement, Modelling and Evalutation of Computer and Communication Systems*, pp. 1–15, 2008.