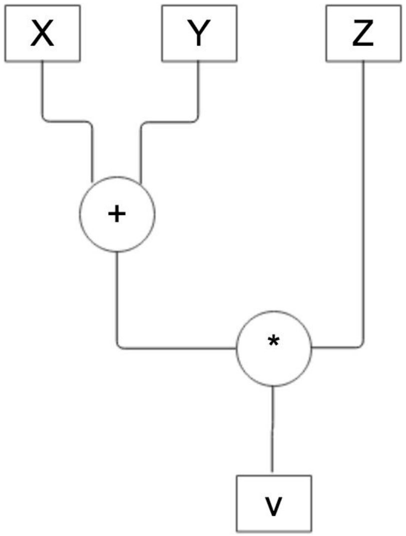
Visualizing Dataflow Graphs of Deep Learning Models in TensorFlow

Kanit Wongsuphasawat, Daniel Smilkov, James Wexler, Jimbo Wilson, Dandelion Mané,
Doug Fritz, Dilip Krishnan, Fernanda B. Viégas, and Martin Wattenberg

Presented by: Halldor Thorhallsson

Background

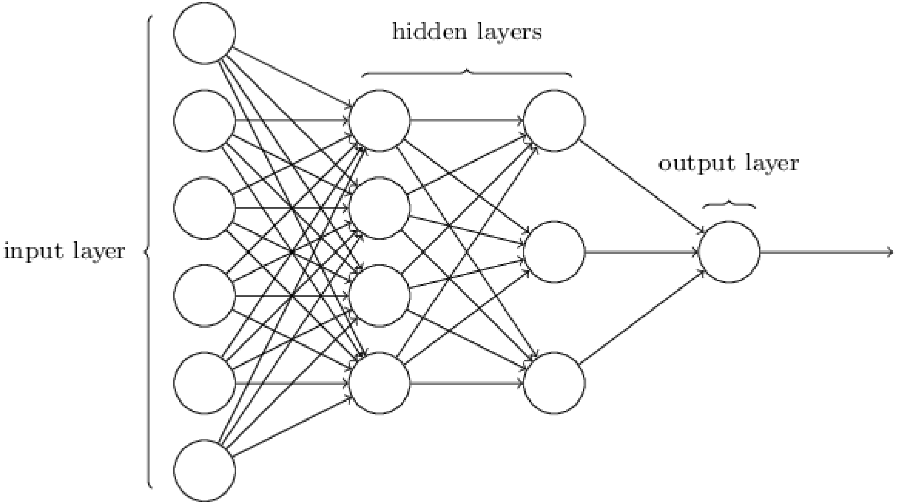
- Dataflow graphs
  - Represent data dependencies between operations
  - **Node**: a unit of computation
  - **Edges**: Data flowing in/out of computation node



V = (X+Y) * Z

Background

- Deep learning models
  - State of the art for many problems in ML
  - Can get very big!
  - Interpretability is a big issue



Artificial neural network. (2017, November 7). In *Wikipedia, The Free Encyclopedia*. Retrieved 06:32, November 13, 2017, from https://en.wikipedia.org/w/index.php?title=Artificial_neural_network&oldid=809235583

TensorFlow

- An API for building neural networks
  - Simplifies the coding process
  - Built around the idea of a dataflow graph
  - Edges are tensors
  - Nodes are mathematical operations

Goals

- Provide an intuitive overview
- Recognize similarities and differences between graph components
- Examine nested structure of a high level component
- Inspect details of low level operations
- Examine quantitative data

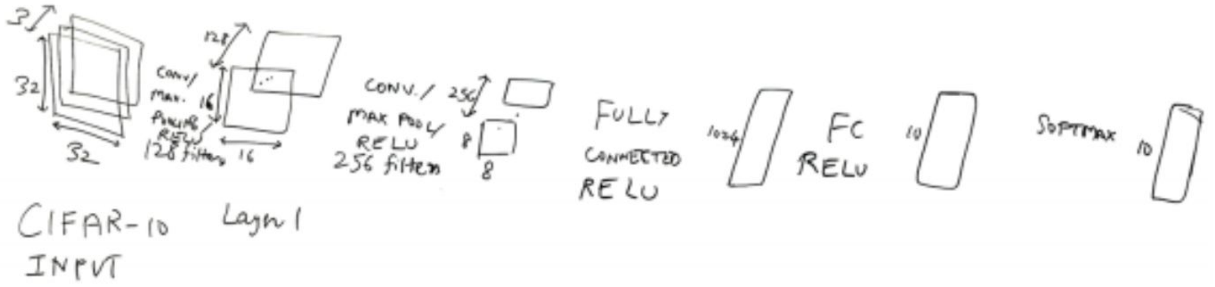
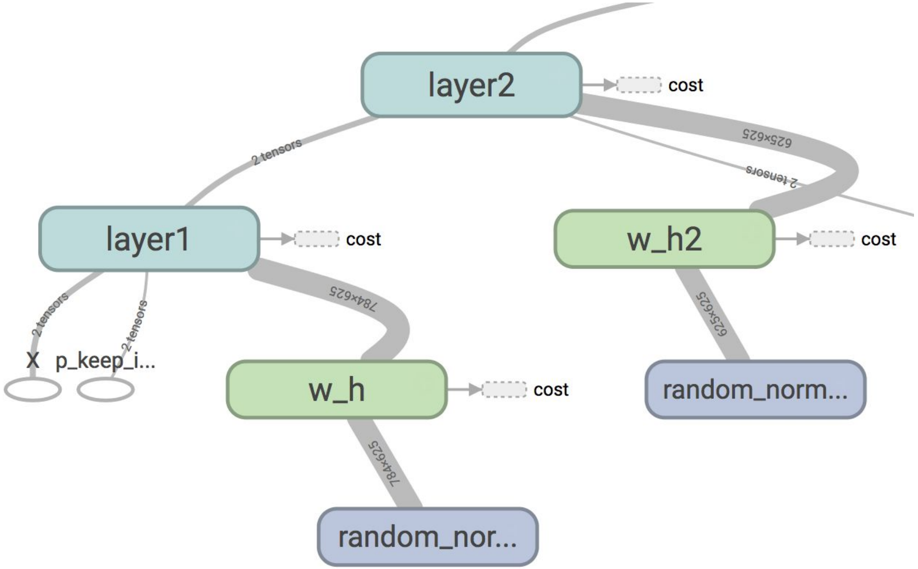


Fig 2. From Visualizing Dataflow Graphs of Deep Learning Models in TensorFlow Kanit Wongsuphasawat et al. *EEE Trans. Visualization & Comp. Graphics (Proc. VAST)*, 2018

Basic Encoding



Problem

- Standard graph layout techniques don’t work well out of the box
  - Mismatch between graph topology and semantics
  - Graph heterogeneity = Not all nodes are created equal
  - Interconnected nodes

Graph Transformations

Remove non-critical nodes

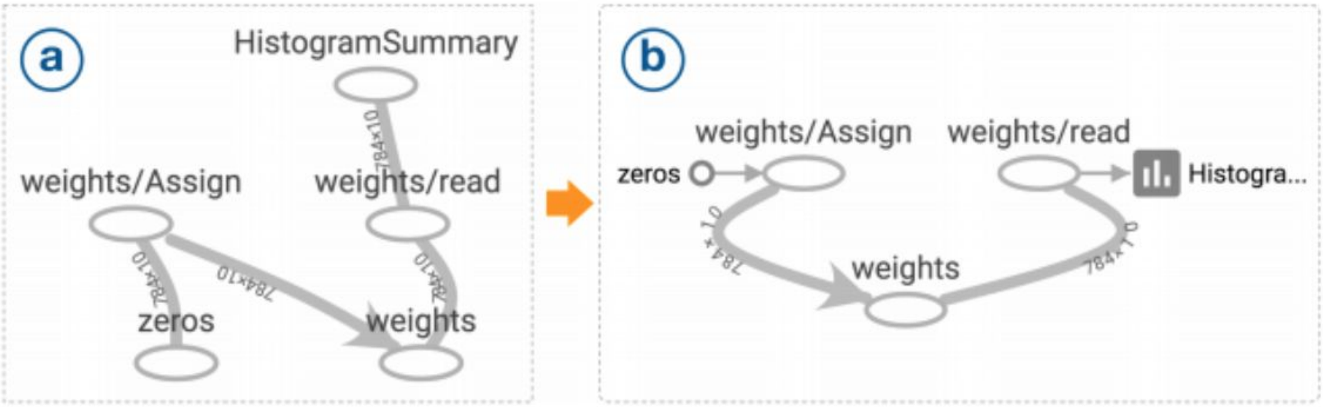


Fig 3. From Visualizing Dataflow Graphs of Deep Learning Models in TensorFlow Kanit Wongsuphasawat et al. *EEE Trans. Visualization & Comp. Graphics (Proc. VAST)*, 2018

Graph Transformations

- Build a hierarchically clustered graph
  - Group operations based on namespaces
  - Bundle edges between groups

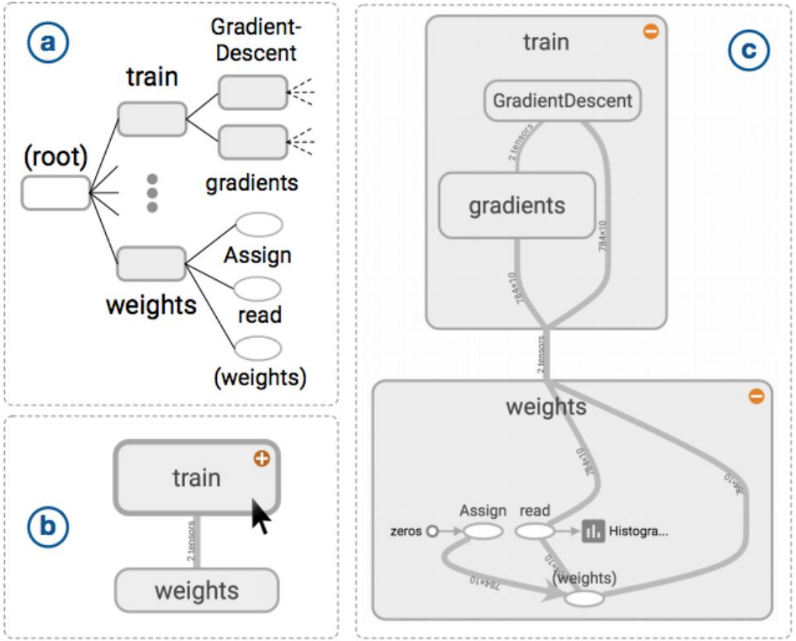
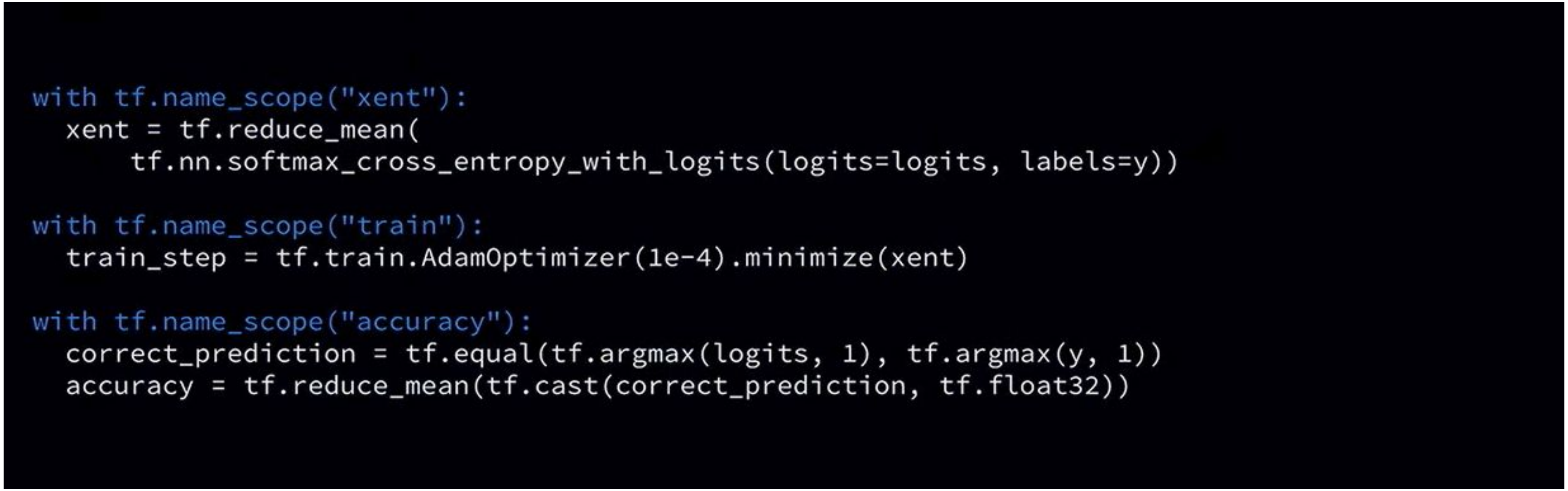


Fig 4. From Visualizing Dataflow Graphs of Deep Learning Models in TensorFlow Kanit Wongsuphasawat *et al. EEE Trans. Visualization & Comp. Graphics (Proc. VAST)*, 2018

Namespaces



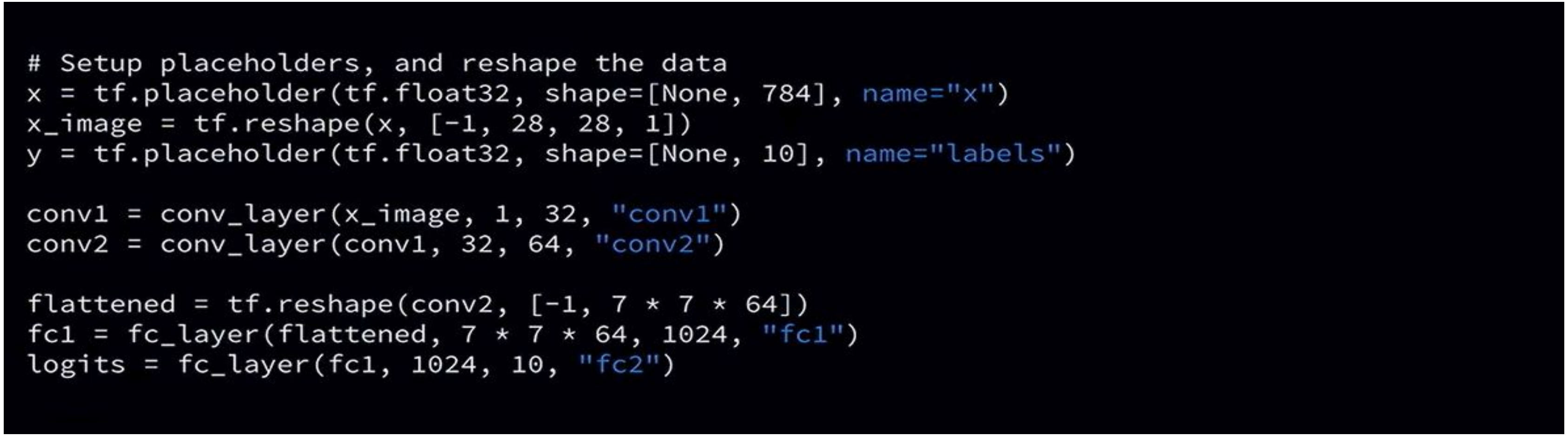
```
with tf.name_scope("xent"):
  xent = tf.reduce_mean(
      tf.nn.softmax_cross_entropy_with_logits(logits=logits, labels=y))

with tf.name_scope("train"):
  train_step = tf.train.AdamOptimizer(1e-4).minimize(xent)

with tf.name_scope("accuracy"):
  correct_prediction = tf.equal(tf.argmax(logits, 1), tf.argmax(y, 1))
  accuracy = tf.reduce_mean(tf.cast(correct_prediction, tf.float32))
```

[Google Developers]. (2017, Feb 15). Hands-on Tensorboard (TensorFlow Dev Summit 2017) [Video File]. Retrieved from https://www.youtube.com/watch?v=eBbEDRsCmv4

Namespaces



```
# Setup placeholders, and reshape the data
x = tf.placeholder(tf.float32, shape=[None, 784], name="x")
x_image = tf.reshape(x, [-1, 28, 28, 1])
y = tf.placeholder(tf.float32, shape=[None, 10], name="labels")

conv1 = conv_layer(x_image, 1, 32, "conv1")
conv2 = conv_layer(conv1, 32, 64, "conv2")

flattened = tf.reshape(conv2, [-1, 7 * 7 * 64])
fc1 = fc_layer(flattened, 7 * 7 * 64, 1024, "fc1")
logits = fc_layer(fc1, 1024, 10, "fc2")
```

[Google Developers]. (2017, Feb 15). Hands-on Tensorboard (TensorFlow Dev Summit 2017) [Video File]. Retrieved from https://www.youtube.com/watch?v=eBbEDRsCmv4

Graph Transformations

Extract Auxiliary nodes from the graph

- Many high-degree nodes turn out to be not that important
  - NoOp nodes
  - Declaring/initializing variables
  - Nodes that compute statistics
- Nodes extracted and put to the right of the whole graph
- Small proxy icons to embedded in neighbouring nodes

Graph Transformations

Deep Learning models have highly repeated structure

- Group nodes with identical subgraphs have the same color
- Finding these subgraphs is NP-hard

VAD Analysis

- **What: Data**
  - Network
- **What: Derived**
  - Cluster hierarchy
  - Edges bundled
  - Nodes removed or embedded
- **Scale:**
  - 1000s of nodes

- **How: Encode**
  - Ellipse/Rectangle marks for nodes, connection marks for links
- **Why: Tasks**
  - Make graph understandable

Demo

Critique

Positive

- Provides an intuitive overview of the dataflow graph
- Makes it easier to debug deep learning models
- Infinitely better than sketches

Negative

- No two-way editing
- Can be hard to tell quickly how the data flows through the network
- Debugging could be improved if integrated with tests
- Diff-viewer would also aid debugging