

Continuous Training with Transfer Learning



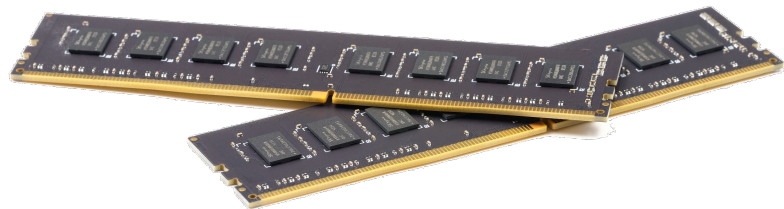
Training Pipeline: **Need Lots of Data**



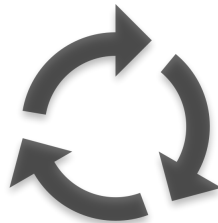
1000 Classes

1000 Images / Class

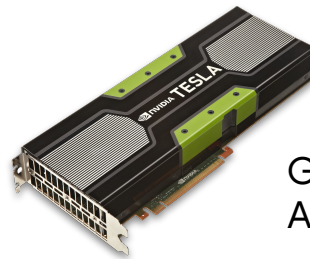
Training Pipeline: Need Compute Resources



Memory



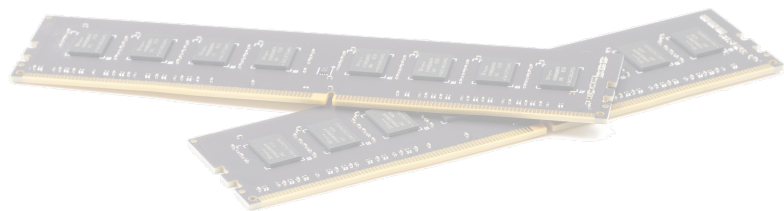
Compute



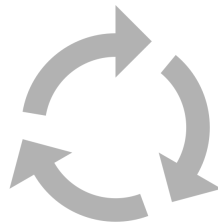
GPU and
Accelerators

Training Pipeline: Need Compute Resources

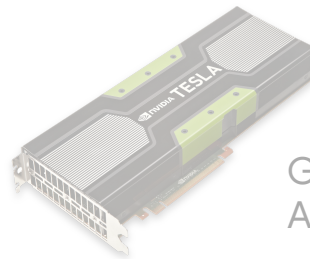
*Computationally Intensive
Repeated Many Times (Epochs)*



Memory



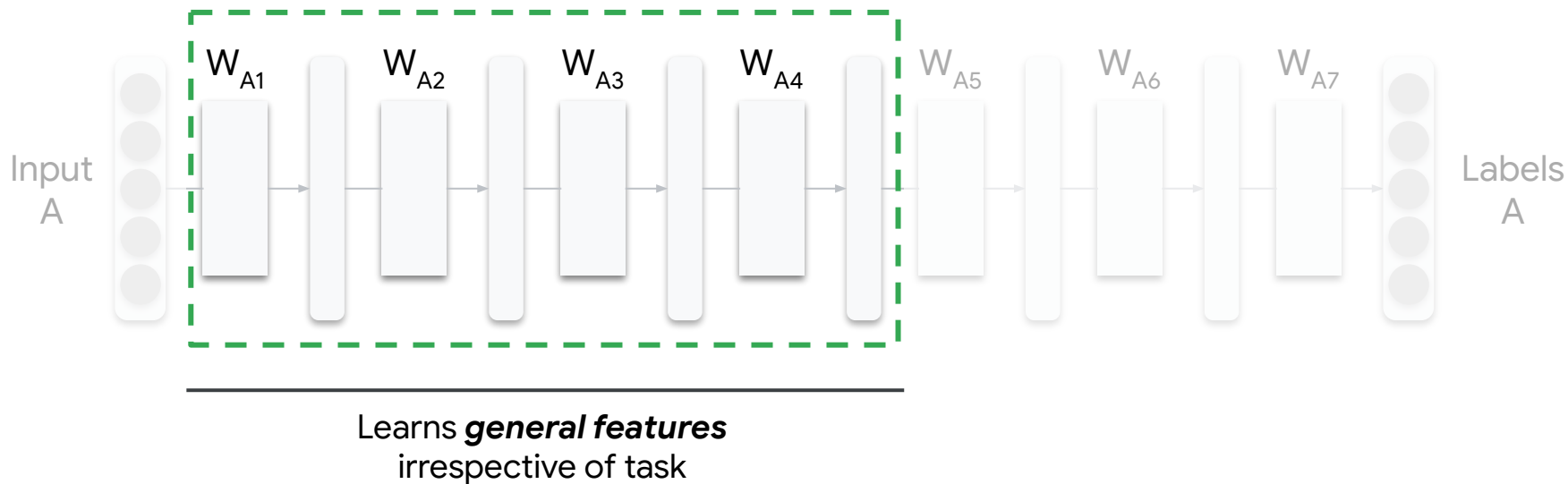
Compute



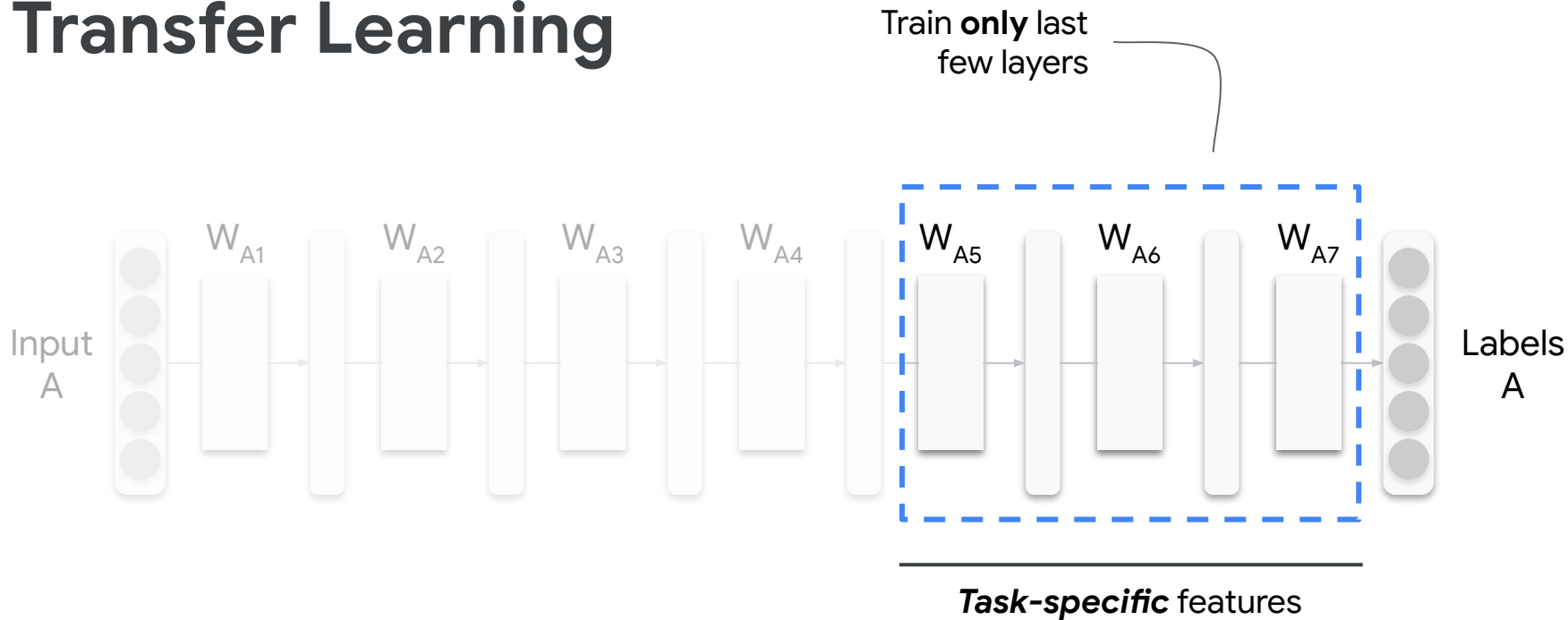
GPU and
Accelerators

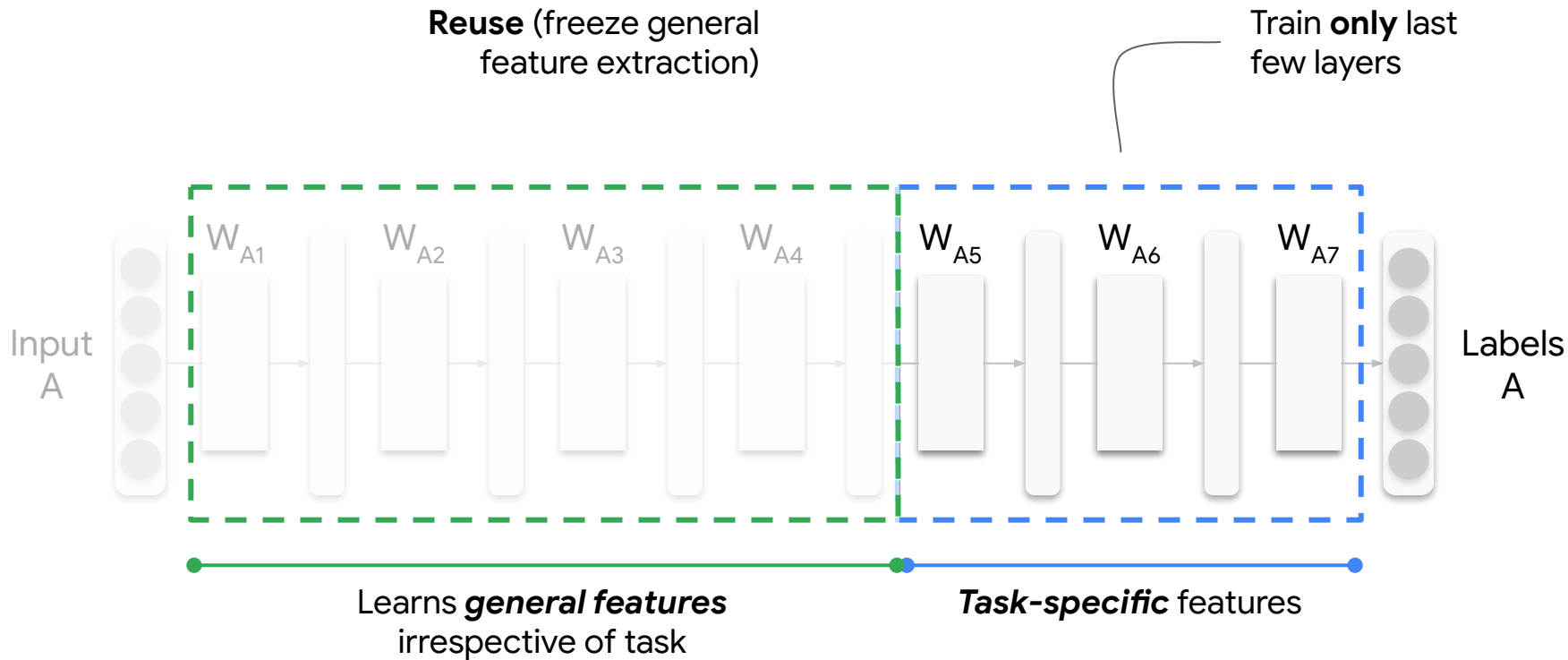
Transfer Learning

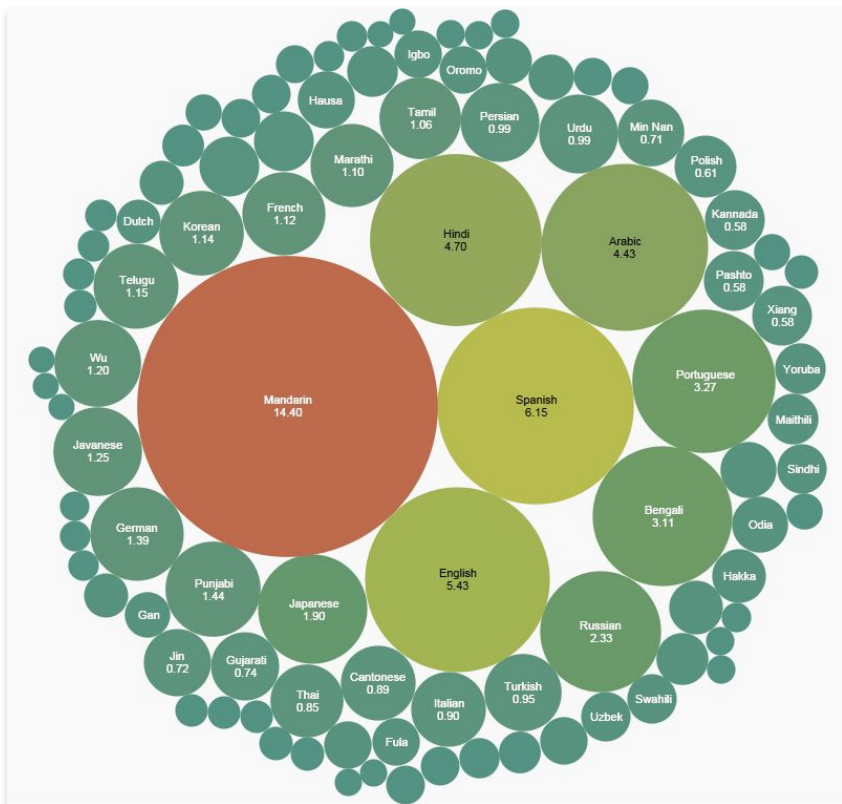
Reuse (freeze general
feature extraction)



Transfer Learning

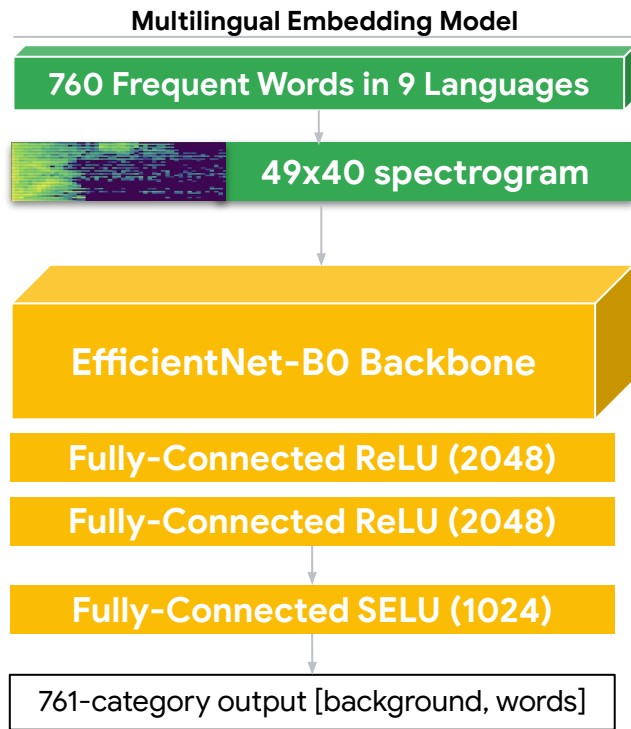






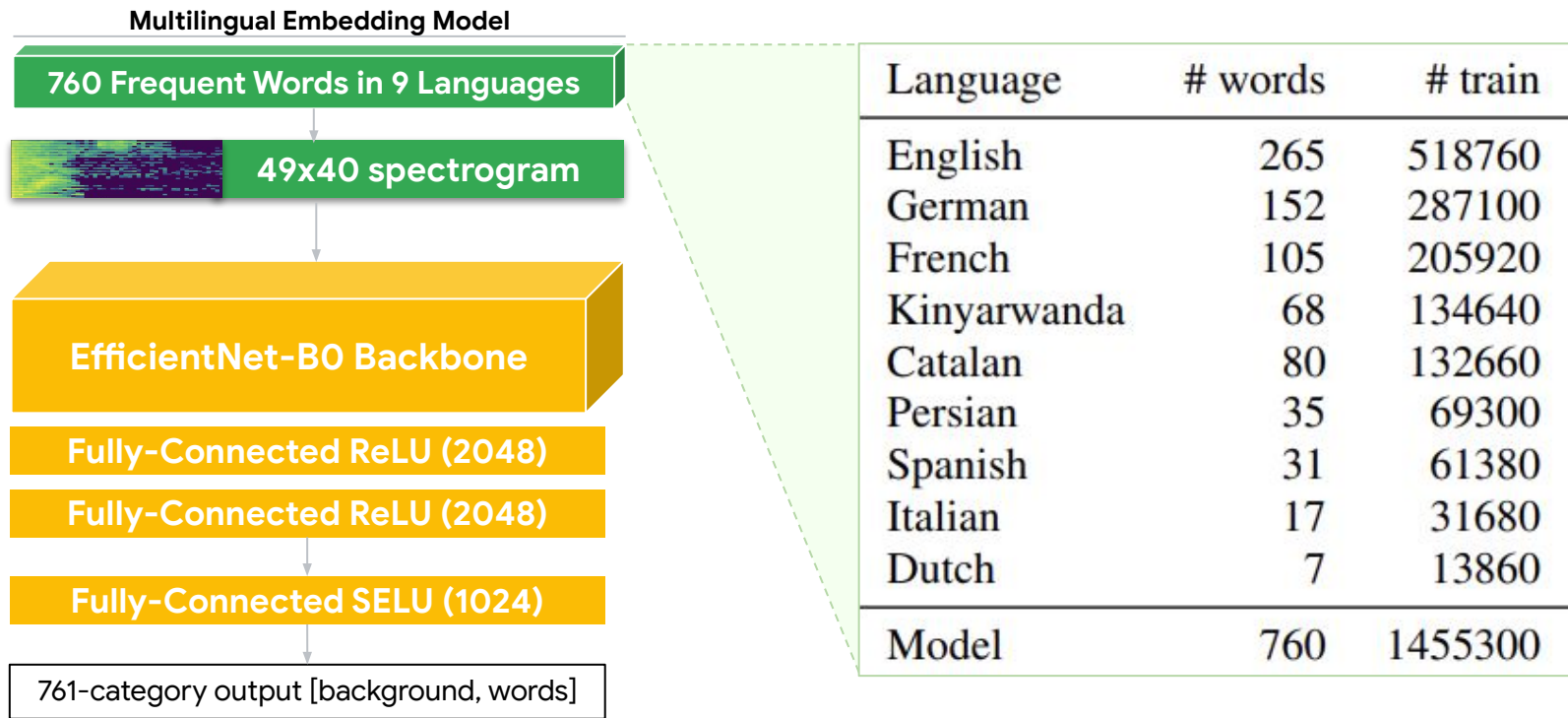
- Speech commands for the **whole planet?**
- For **more than** just voice assistants

Multilingual Embedding Model

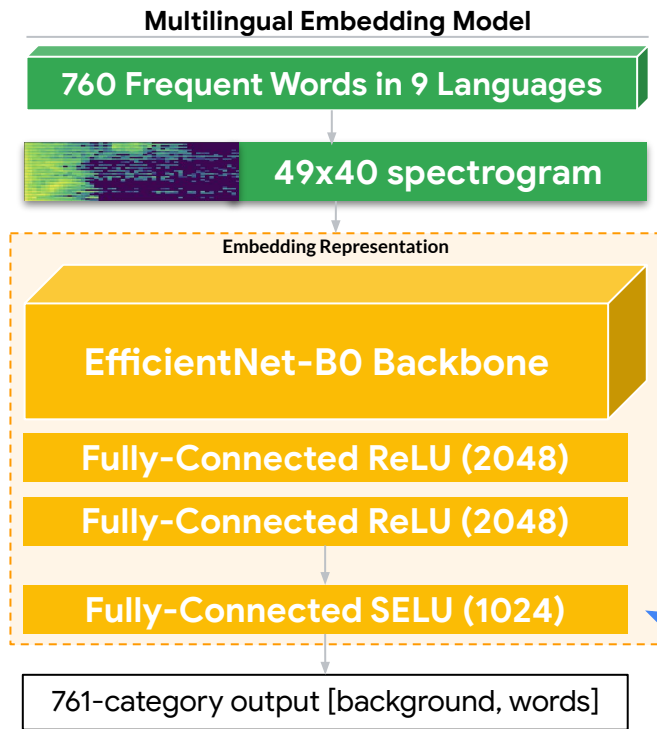


Simple classifier for 760 frequent words in 9 languages

Multilingual Embedding Model

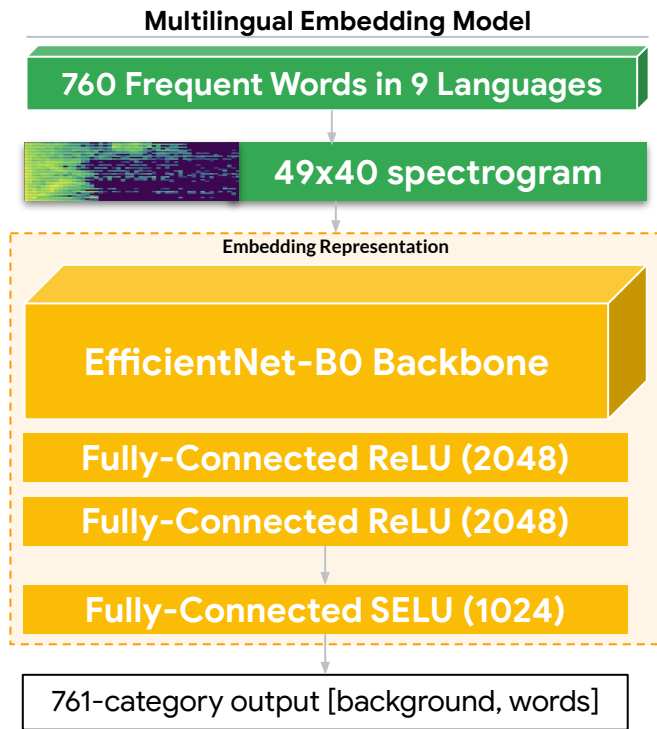


Few-Shot Keyword Spotting

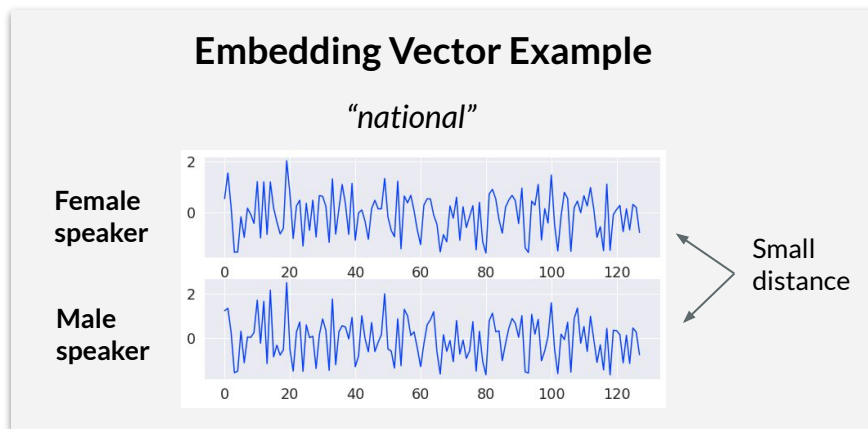


Use **penultimate layer** for embedding representation in **keyword-spotting model**

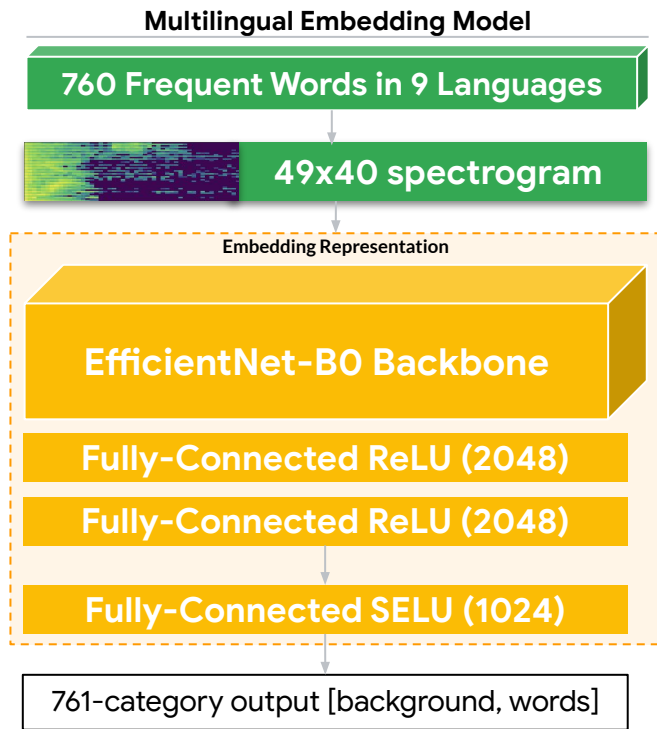
Multilingual Embedding Model



Penultimate layer **output** as embedding vector

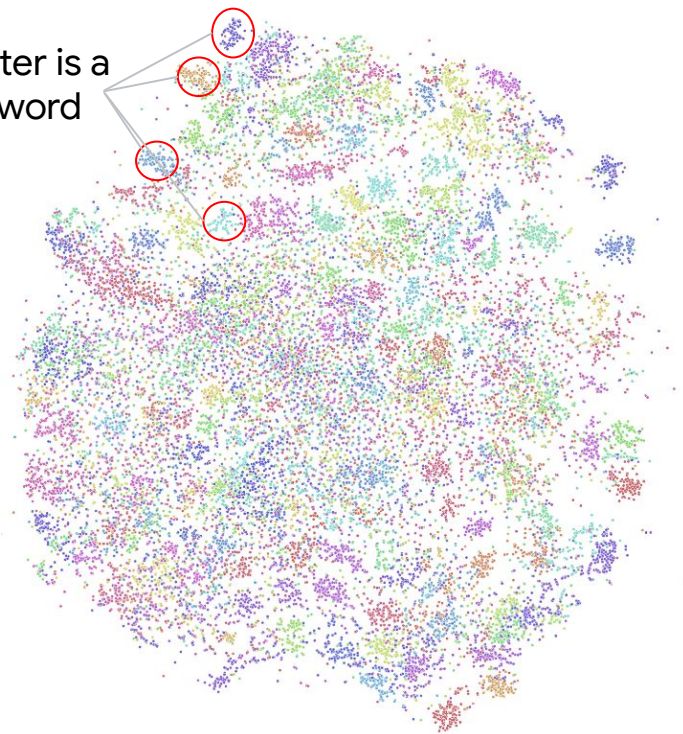


Multilingual Embedding Model

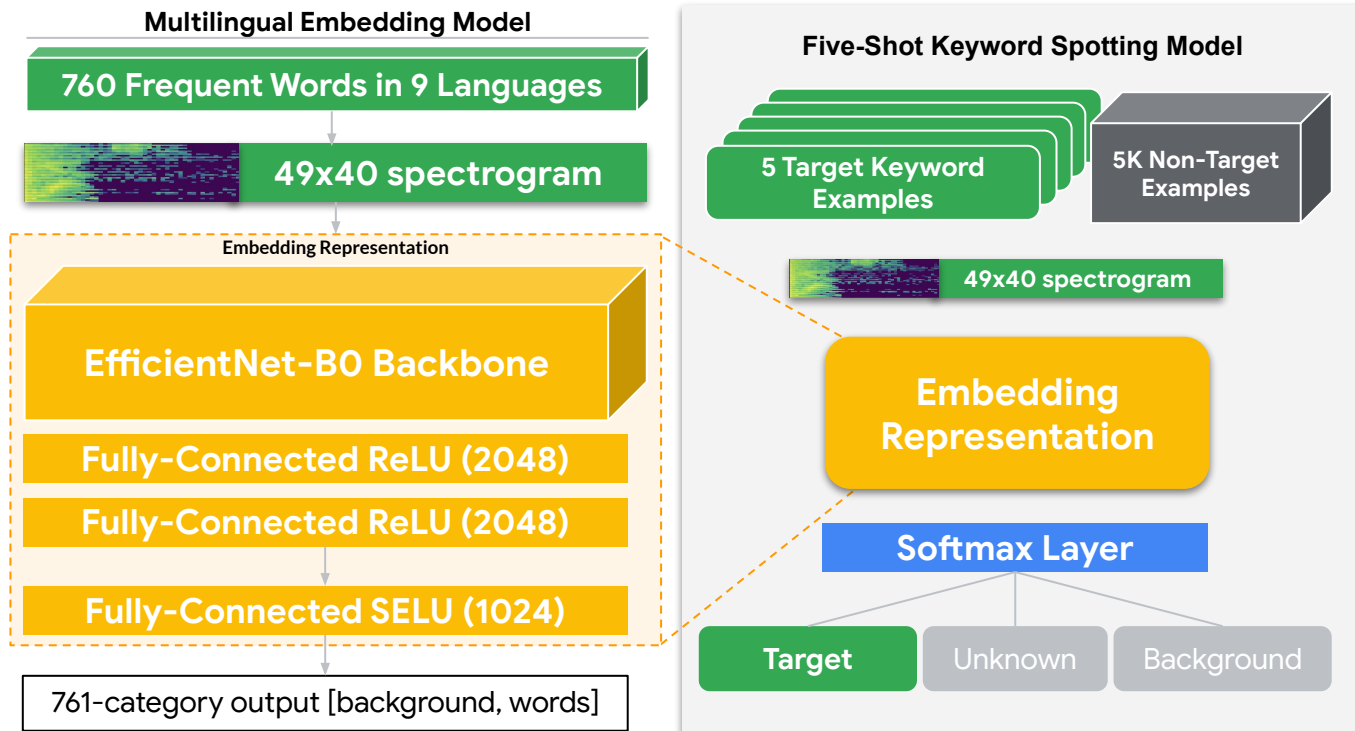


t-SNE view of 165 novel words (not used to train the classifier)

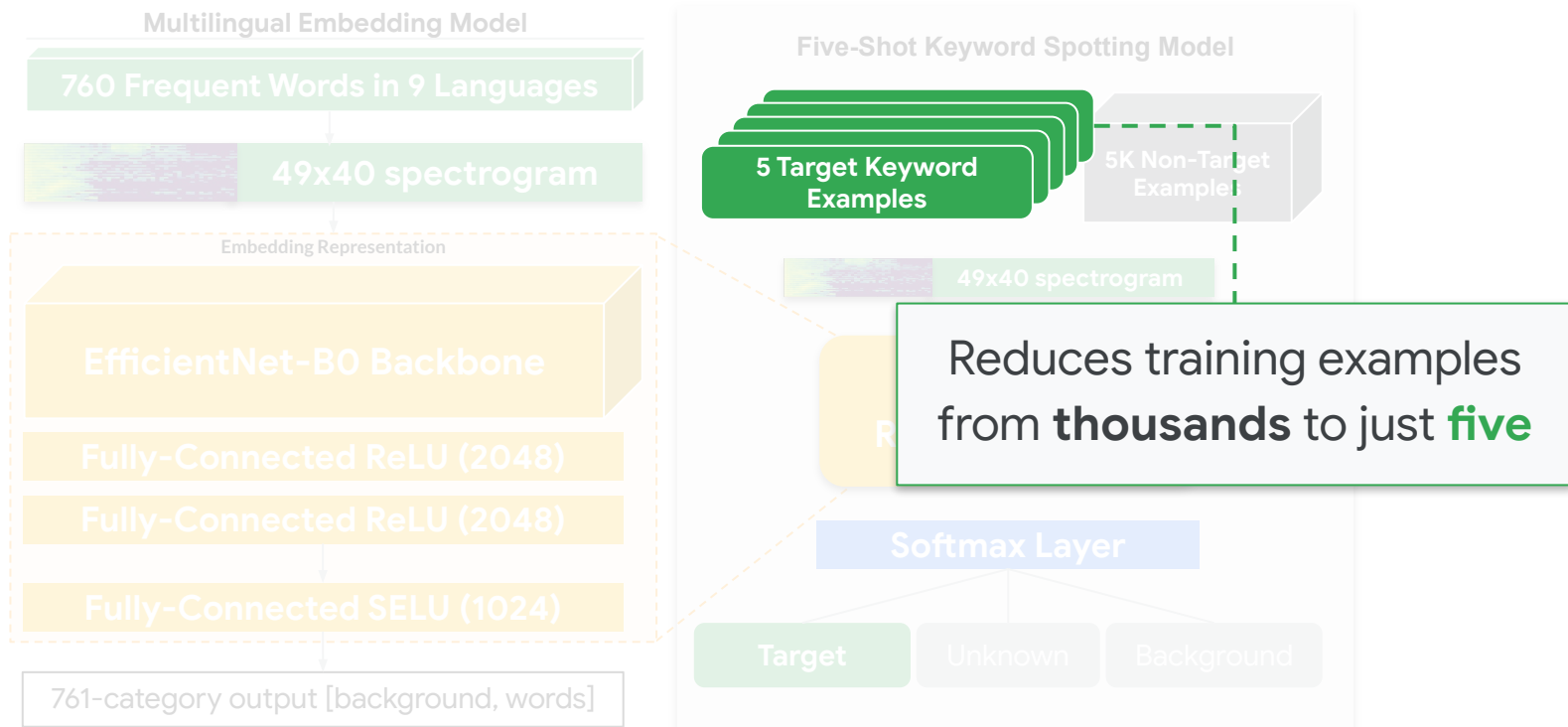
Each cluster is a different word



Few-Shot Keyword Spotting



Few-Shot Keyword Spotting



Transfer Learning

