# 理论介绍

## 应用场景

数据量少，网络模型复杂，难以充分训练，导致模型效果不佳的情况。

最先应用于图像处理问题中。

## 思想

1. 在复杂的深度神经网络，比如CNN中，底层的网络主要用于提取基础的特征，比如图片中的点线特征，如图1.1。这些底层的网络与具体任务无关，具有普适性。

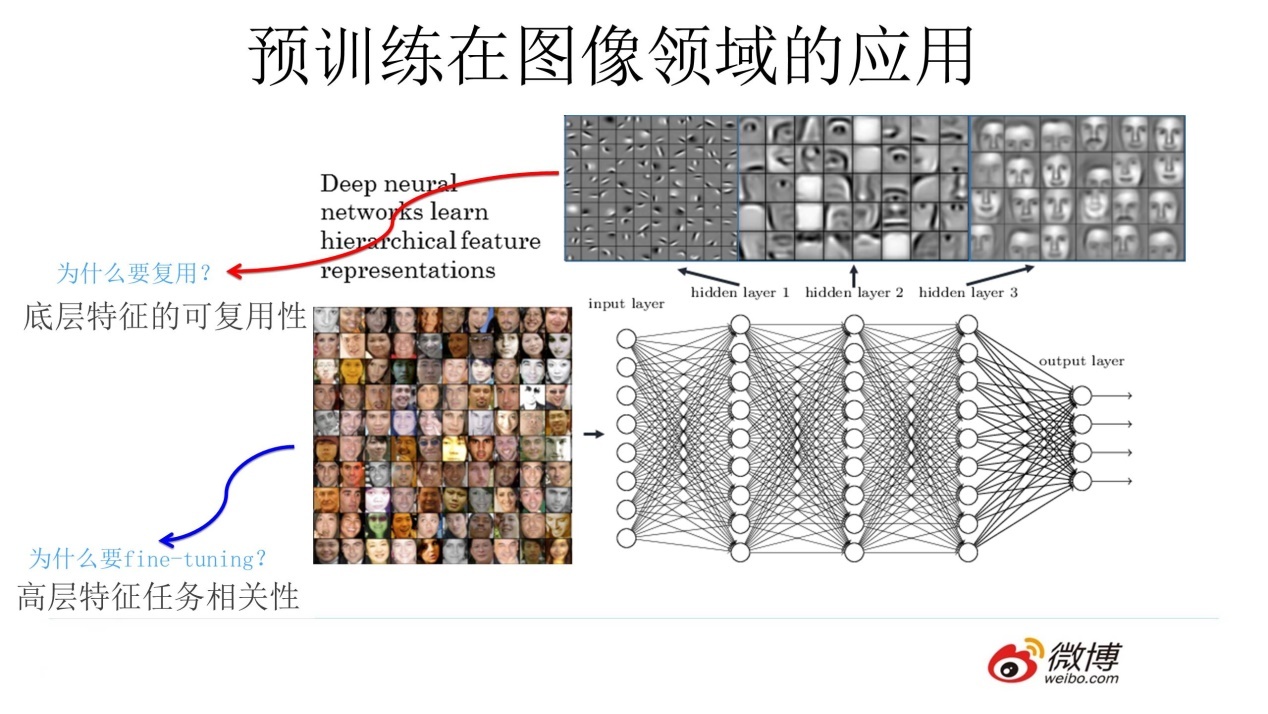


图 1.1网络结构

1. 因此，可以将复杂网络的前几层抽出，单独构成一个简单模型，然后使用比较类似的大量数据进行训练。
2. 然后将训练好的几层参数直接赋给复杂网络的前几层，在训练复杂网络时，前几层的参数可以一直保持不变，类似于Word Embedding这种方式，叫做Frozen；也可以在训练时跟着一起优化，叫做Fine-Tuning。如图1.2所示。

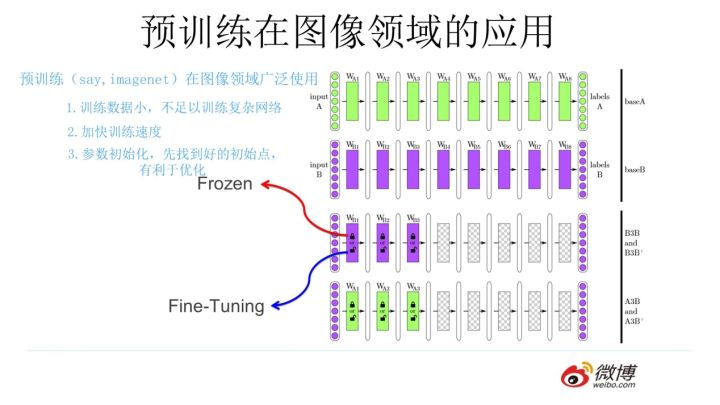


图 1.2预训练

## 在NLP领域的应用

Word Embedding、ELMO、GPT等