Transfer learning

To leverage ....

Source: https://towardsdatascience.com/a-comprehensive-hands-on-guide-to-transfer-learning-with-real-world-applications-in-deep-learning-212bf3b2f27a

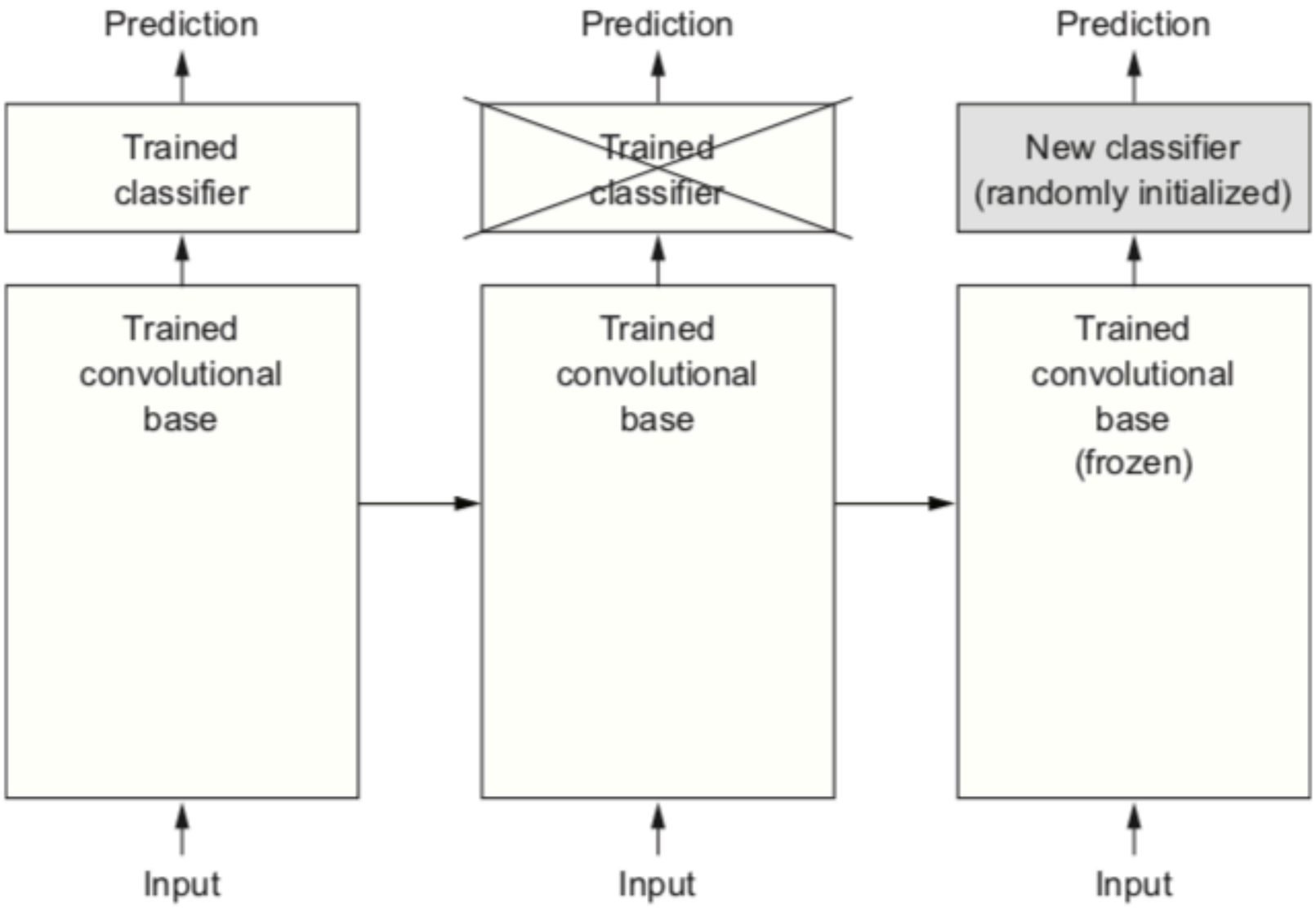
- A common and highly effective approach on small image datasets
- Use a pre-trained network that was previously trained on a large-scale dataset
- If the original dataset is large and general enough, the spatial hierarchy of features learned by the pretrained network can act as a generic model of visual world
- Thus, the way the net extracts features is useful to many different visiom problem

Source: Deep learning with Python by Francois Chollet

Transfer learning

Feature extraction

- Two approaches to use pre-trained network: feature extraction and fine-tuning

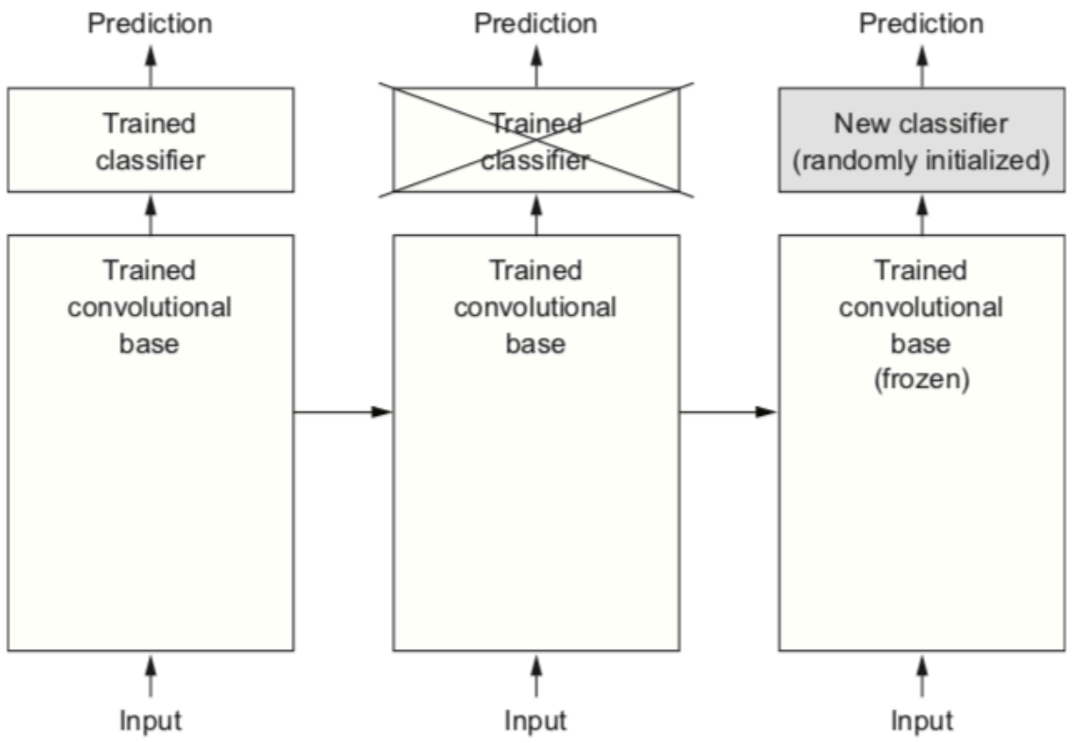


Source: Deep learning with Python by Francois Chollet

Transfer learning

Feature extraction

- Can we re-use the classifier?
- Representations learned by convnet base likely more generic, thus reusable
- Representations learned by classifier more specific to the set of classes the model was arranged to be trained on, not so reusable



Source: Deep learning with Python by Francois Chollet

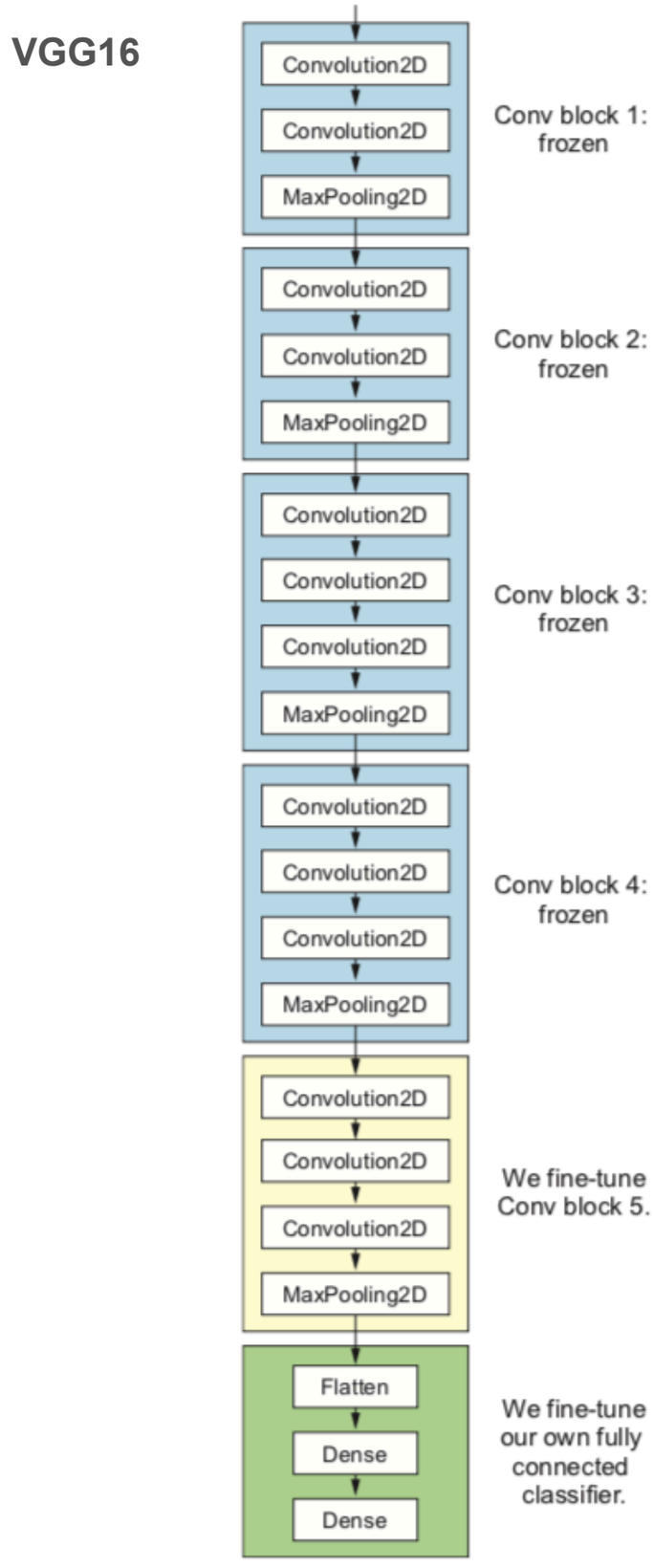
Transfer learning

Fine tuning

- **After we have trained the added classifier with the convnet base frozen**, we can unfreeze a few top layers of the frozen base model (layers that are near to the classifier)

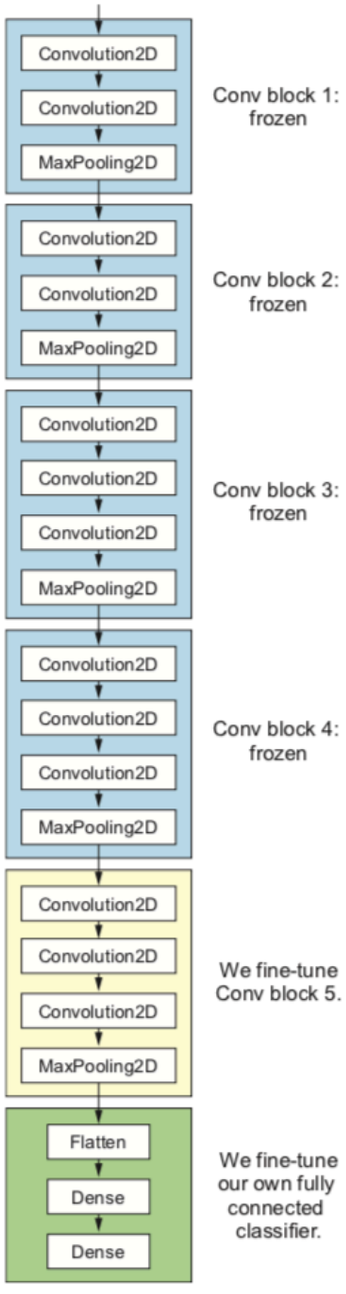
- Fine-tuning: we train the newly unfreezed layers together with the classifier

Source: Deep learning with Python by Francois Chollet



Transfer learning

Feature extraction

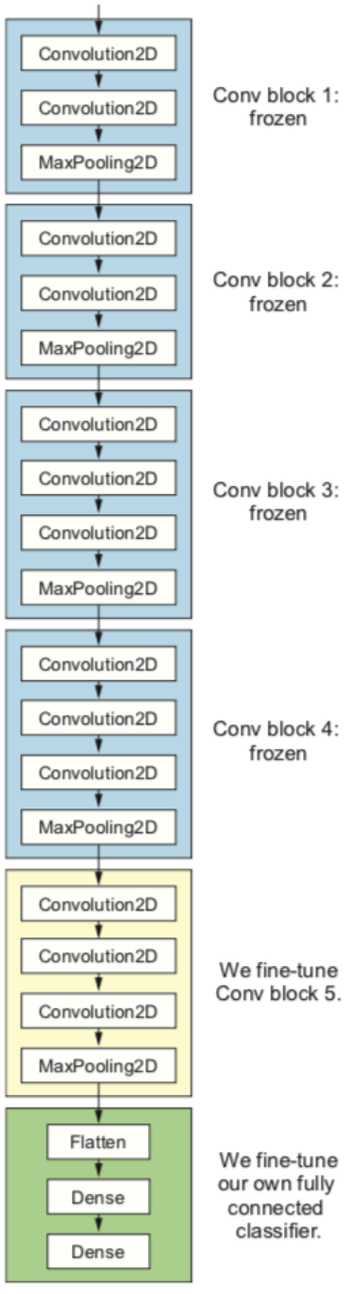


Source: Deep learning with Python by Francois Chollet

- Must the classifier be trained before we perform a fine-tuning?
- Yes, else the errors back-propagated will be too large and distort the weights of the just unfreezed layers
- Note: we are doing **fine-tuning** on both the unfreezed layers and classifier, **not re-train** the both

Transfer learning

Feature extraction



Source: Deep learning with Python by Francois Chollet

- Can unfreeze and fine-tune more layers?
- Things to consider: Earlier layers in the base net (layers that are near to input) encode more generic and more reusable features; this is something you want to keep
- Later layers are more specific, you want to fine-tune them to make them fit for your problem
- Pre-trained networks usually have much more parameters and are much more powerful (that's why you want to use them); re-train more layers will easily lead to overfitting on small dataset