Transferência de Aprendizado é a técnica em que se utiliza um conjunto de parâmetros aprendidos a partir de um determinado domínio D e tarefa T, e se usa esse aprendizado para aplicar num outro Domínio D' e/ou tarefa T'.  
https://core.ac.uk/download/pdf/81905331.pdf  
  
No cenário de redes profundas, em que é preciso uma grande quantidade de dados e épocas para garantir que houve convergência, a transferência de aprendizado é amplamente utilizada. Em particular para dados não estruturados pois esses dados possuem certas características que são independentes da base de dados. Vou dar exemplos com as áreas de texto e imagens.  
  
- No caso de texto, uma representação bem aprendida da relação entre palavras em uma frase num determinado contexto pode ser muito útil em outros cenários. Por exemplo, assuma que foi treinada uma rede neural utilizando textos da Wikipedia em português com o objetivo de classificar os artigos em 200 categorias diferentes. A última camada da rede neural é a responsável por fazer a classificação, mas todas as outras trabalharam no sentido de obter, de forma hierarquica, uma \*representação\* do texto que seja útil para classificação. Se o treinamento foi bem sucedido então o que a rede está fazendo é extrair boas características do texto que permitem diferenciar artigos em 200 categorias diferentes. Nesse ao usar transferência utilizamos o modelo base, removendo a última camada (especializada nas 200 categorias) e usar a representação obtida da penúltima camada para uma outra tarefa, por exemplo classificação de sentimentos (positivo, negativo, neutro) em postagens de redes sociais. A premissa aqui é que as mesmas características obtidas das combinações de palavras e frases que é útil para diferenciar artigos, também é para classificar sentimentos. Aqui o domínio é diferente (D = artigos wiki, D' = posts de redes sociais) e a tarefa é diferente (T=classificaçao em categorias, T'=classificacao de sentimentos)  
\*\* Note que, quando usamos um método de extração de características (ex. TF-IDF) o que estamos fazendo é confiar que essa estatística textual será suficiente para alguma tarefa. No fundo, ao usarmos uma rede neural pré-treinada estamos utilizando uma premissa similar.  
  
- No caso de imagens, as convoluções aprendidas pela rede neural para uma base de dados grande extraem/filtram características de regiões da imagem, obtendo informações sobre cores, bordas, texturas e outros. Considerando que estamos trabalhando com imagens RGB, toda imagem potencialmente possui características de: cor, textura, forma, etc. Como a base de dados imagenet é grande suficiente e tem variedade tanto de exemplos quanto de categorias, considera-se que essas características também são muito diversas. O fato da representação aprendida ser boa o suficiente para classificar imagens em 1000 categorias, permite que essa também seja útil em outros cenários. Notar que a ImageNet possui classes relacionadas a plantas, árvores e outros, o que nesse caso ajudou. Portanto aqui o domínio é similar (D = imagens fotográficas, incluindo plantas, D' = imagens de plantas, de forma que D' é um subconjunto de D) e a tarefa é diferente (T=classificaçao em 1000 categorias, T'=classificacao de doenças, podemos considerar que T' é de granularidade fina em comparação com T).  
\*\* Como eu expliquei na aula, caso o domínio de interesse não esteja contido na base original (ex. ImageNet), uma mera extração de características não seria suficiente e então teremos que necessariamente recorrer a um ajuste fino, o que exige maior quantidade de imagens.

Eis uma "visualização" dos filtros aprendidos nas primeiras camadas convolucionais de redes neurais.  Eles são (parte de cima) detectores de bordas, e (parte de baixo) detectores de regiões com certos tipos de cores e texturas). É muito comum que esses filtros/kerneis convolucionais das primeiras camadas sejam muito parecidos independente da base de dados que se utiliza como entrada, pois são consideradas características de pixel de baixo nível. Há estudos ainda que relacionam as primeiras camadas convolucionais a bancos de filtros de Gabor por exemplo.

