O código fornecido para a utilização da CNN pre-treinada, difere em alguns pontos importantes do código construído do zero. As principais diferenças são

- **Definição do modelo e pré-treinamento:**No código fornecido o modelo MobileNetV3\_small é previamente pre-treinado e carregado no ImageNet, o que fornece uma vantagem pois começa com as features já aprendidas do grande dataset. Já se fossemos definir um dataset do zero teríamos que inicializar todos os pesos de forma randômica e criar a arquitetura da CNN do zero, o que normalmente precisaria de mais épocas e dados para o modelo aprender as features

- **Modificação de camada e melhor otimização:**Já que o MobileNet é construído para as 1000 classes do ImageNet, esse código modifica a ultima camada para fornecer 10 classes do  dataset FashionMNIST. Enquanto se fosse um modelo do zero, todas as camadas incluindo a de saída seriam customizadas para uma tarefa específica.

- **Congelamento de camadas:**Utilizando a transferência de aprendizado, todas as camadas com exceção da última camada são congeladas, o que significa que os pesos dessas camadas não serão atualizados, o que resulta na diminuição de processamento computacional.

- **Velocidade de treinamento e convergência**: Usando a transferência de conhecimento existe uma tendencia do modelo convergir mais rápido, pois o modelo já tem um base solida de features, o que resulta em menos épocas para atingir uma boa acurácia.

**- Configurações de otimização**: nesse código as otimizações apenas precisam atualizar os parâmetros da última camada devido ao congelamento das outras, reduzindo o esforço computacional e o tempo gasto.