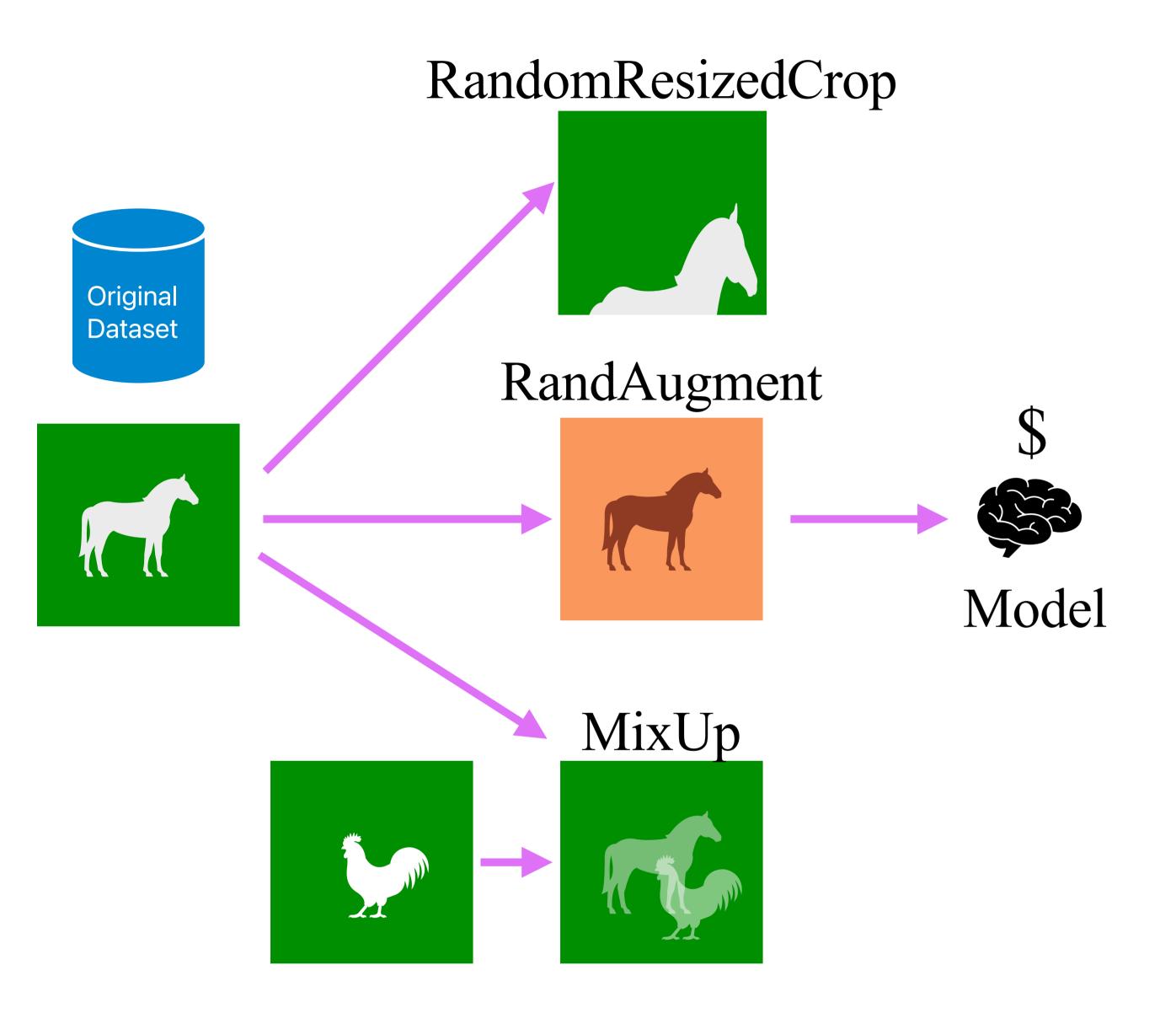
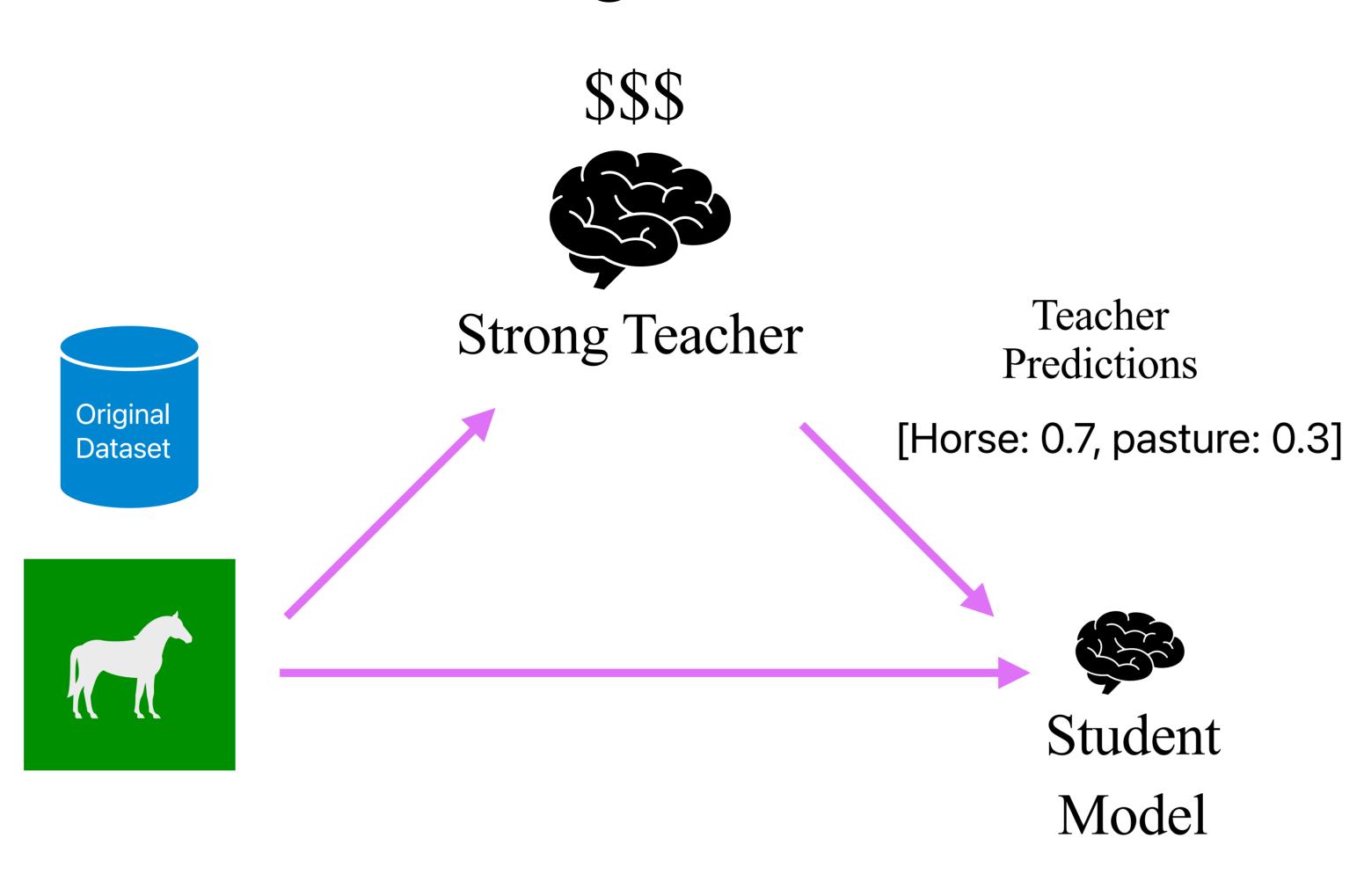
## Data Augmentation



- •Model-specific code change
- Specialized for every new model
- •Sub-optimal accuracy vs Knowledge Distillation Code:

dataset = ImageNetDataset(train\_dir)
data\_loader = DataLoader(dataset, transforms)

# Knowledge Distillation

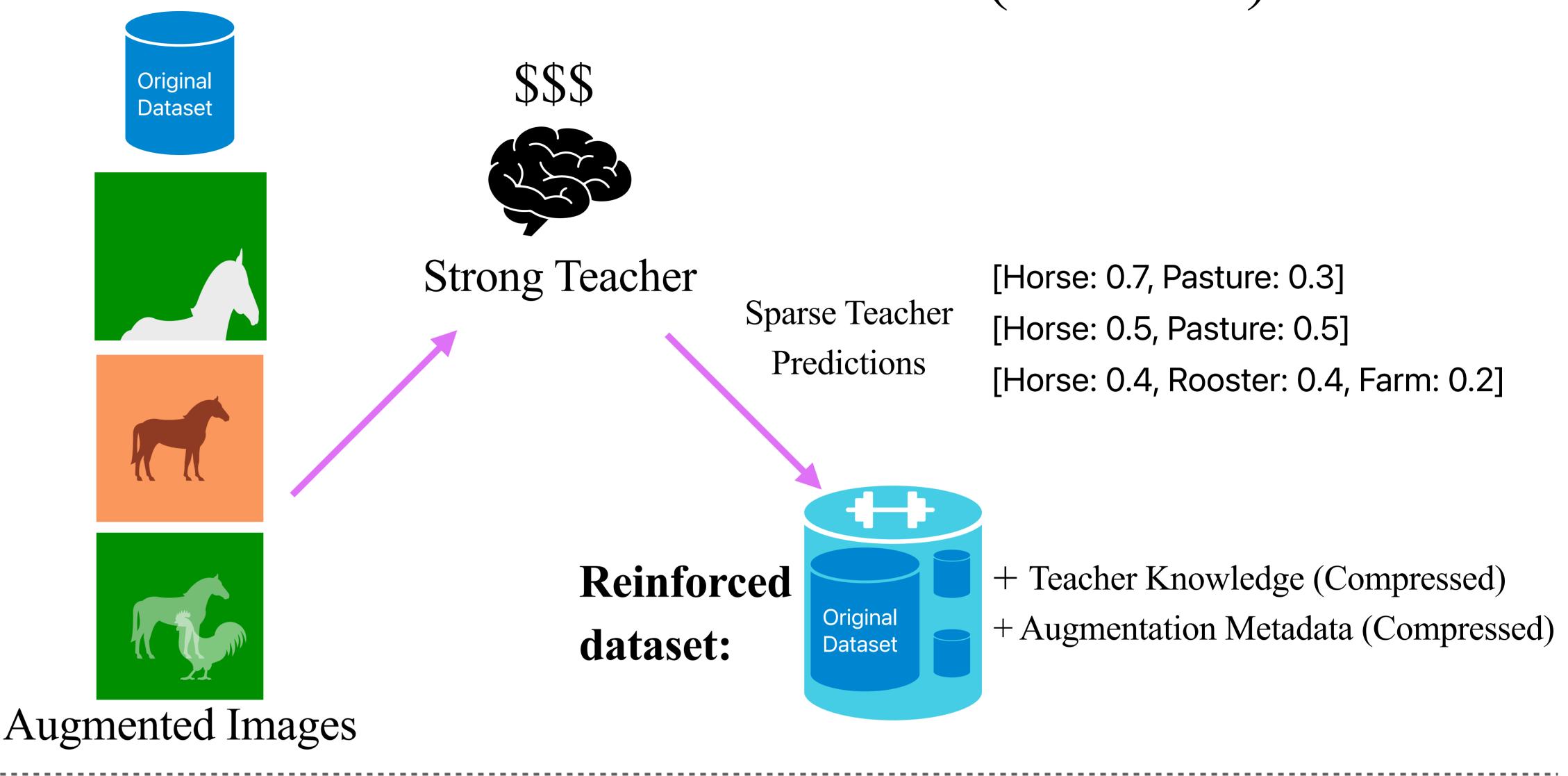


- •Significant overhead for every new model
- •Significant code change and hyperparameter tuning

#### Code:

model\_output = model(input) teacher\_output = teacher(input)

## Dataset Reinforcement (one time)



### Training with Reinforced dataset (several times)

No new hyperparameters

•Minimal code change:

dataset = ImageNetDataset(train\_dir)
data\_loader = DataLoader(dataset, transforms)

dataset = ImageNetPlusDataset(train\_dir)
data\_loader=DataLoader(dataset, basic\_transforms)

