CS230
Healthcare Case Study

Kian Katanforoosh
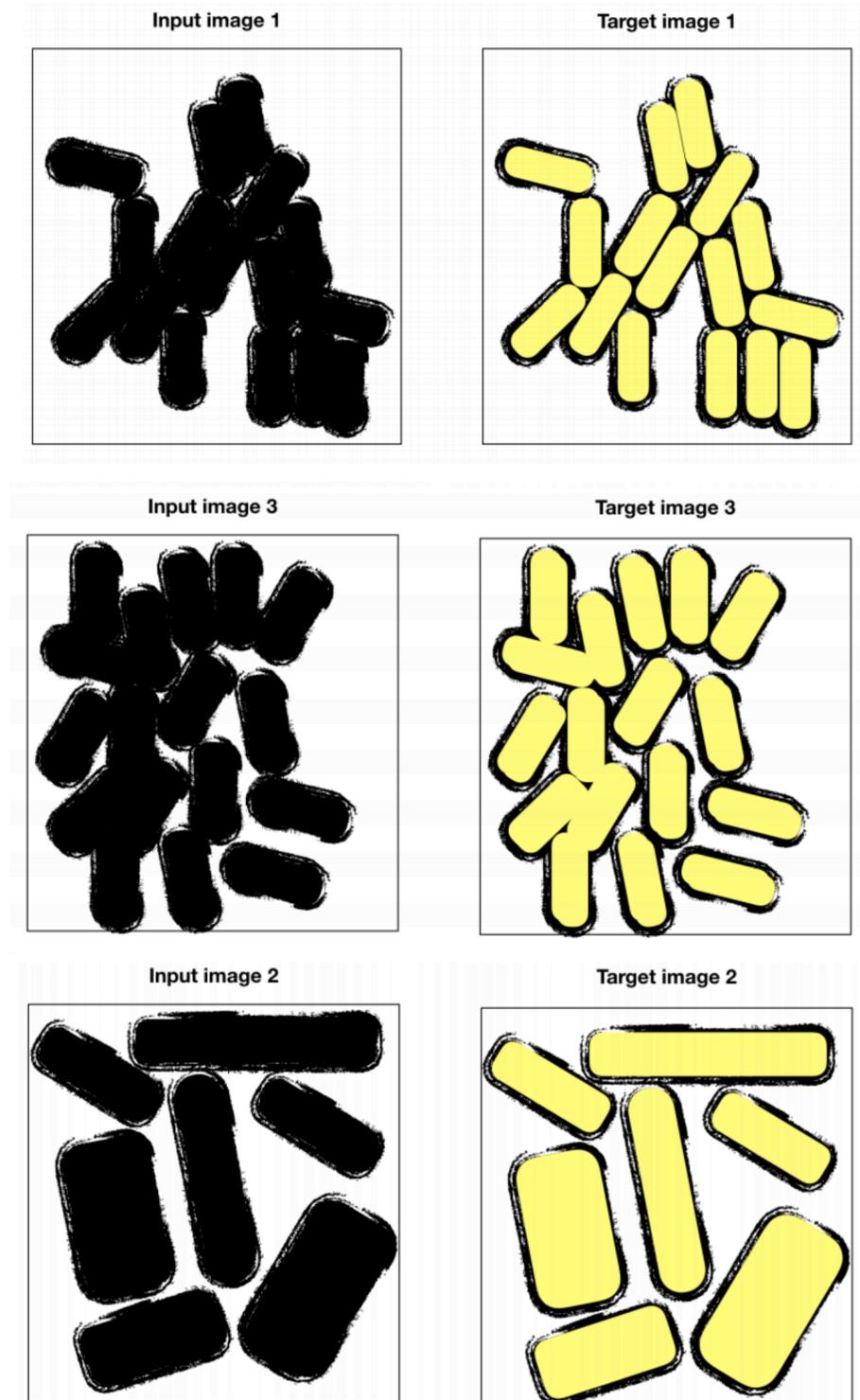
Problem statement: cell segmentation

Goal: Determine which parts of a microscope image corresponds to which individual cells.

Data: Doctors have collected 100,000 images from microscopes and gave them to you. Images have been taken from three types of microscopes:

Type A	50,000 images
Type B	25,000 images
Type C	25,000 images

Question: The doctors who hired you would like to use your algorithm on images from microscope C. How you would split this dataset into train, dev and test sets?

Input image 1   Target image 1
Input image 3   Target image 3
Input image 2   Target image 2

Kian Katanforoosh

Question: The doctors who hired you would like to use your algorithm on images from microscope C. How you would split this dataset into train, dev and test sets?

Answer:
i) Split has to be roughly 90,5,5. Not 60,20,20.
ii) Distribution of dev and test set have to be the same (contain images from C).
iii) There should be C images in the training as well, more than in the test/dev set.

Question: Can you augment this dataset? If yes, give only 3 distinct methods you would use. If no, explain why (give only 2 reasons).

Answer: Many augmentation methods would work in this case:
– cropping
– adding random noise
– changing contrast, blurring.
– flip
– rotate

Input image

Expected output

First try: You have coded your neural network (model M1) and have trained it for 1000 epochs. It doesn't perform well.

Model

Transfer Learning: One of your friends suggested to use transfer learning using **another labeled dataset** made of 1,000,000 microscope images for skin disease classification (very similar images).

A model (M2) has been trained on this dataset on a 10-class classification. Here is an example of input/output of the model M2.

Input image

Expected output

Model
(20 fully-connected layers)   softmax

0.32	Probability of disease 1
0.08	Probability of disease 2
0.04	Probability of disease 3
0.06	Probability of disease 4
0.36	Probability of disease 5
0.01	Probability of disease 6
0.02	Probability of disease 7
0.01	Probability of disease 8
0.01	Probability of disease 9
0.09	Probability of disease 10

Question: You perform transfer learning from M2 to M1, what are the new hyperparameters that you'll have to tune?

Kian Katanforoosh
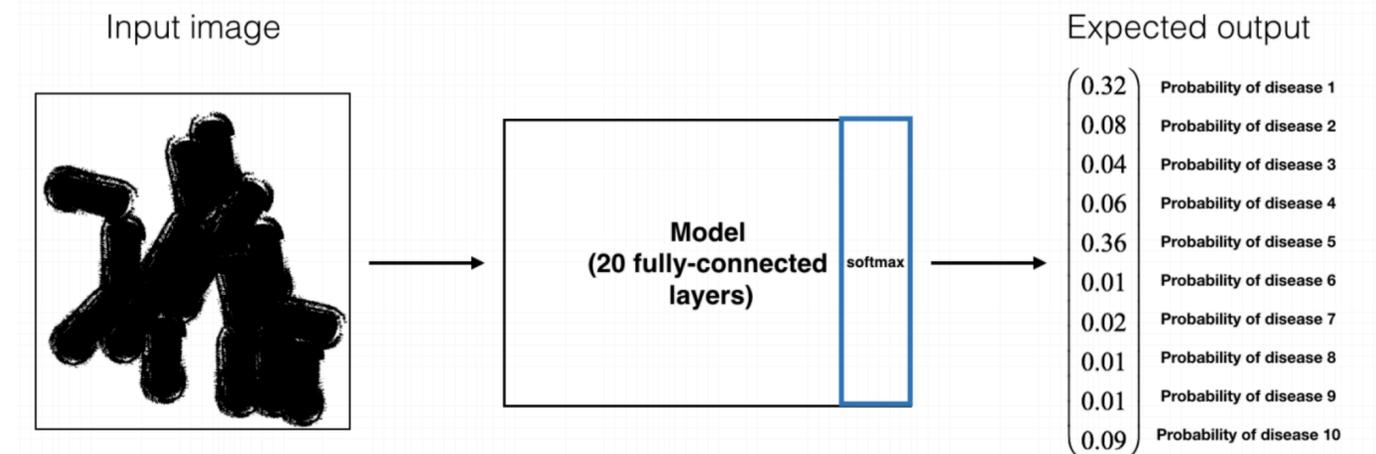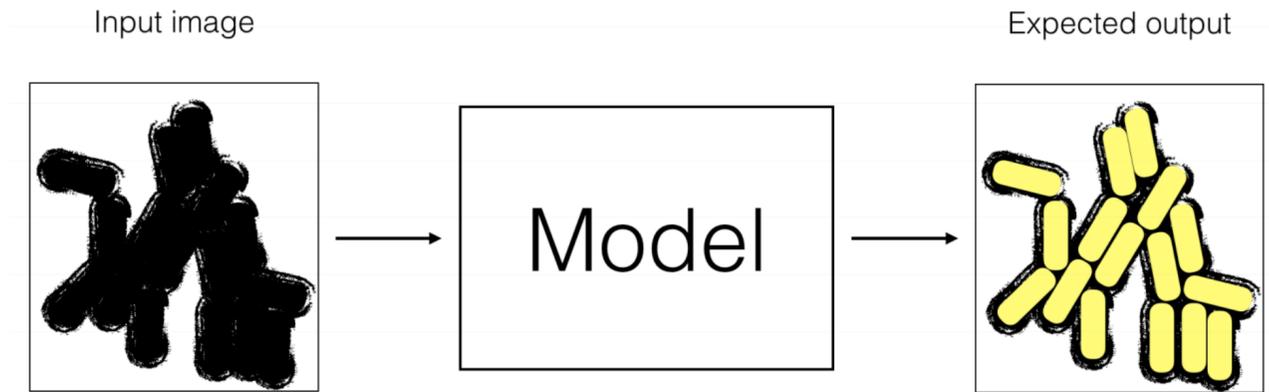
Question: You perform transfer learning from M2 to M1, what are the new hyperparameters that you'll have to tune?
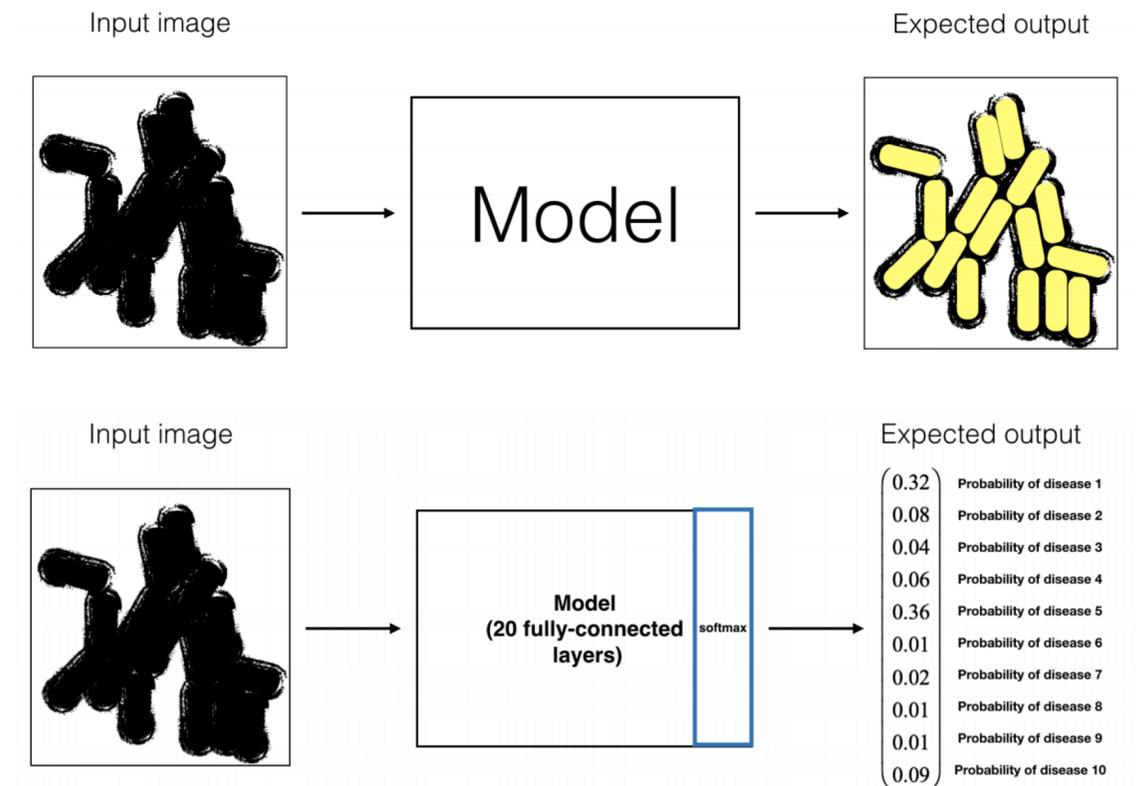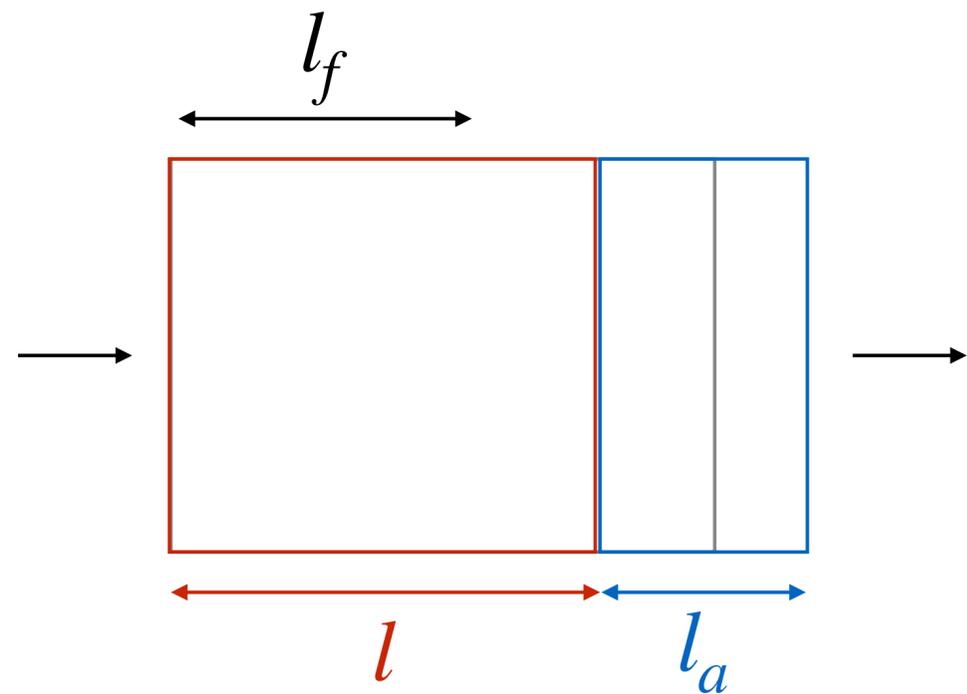
l = number of layers transferred from M2

l_a = number of new layers added to the new model's head

l_f = number of frozen layers

Kian Katanforoosh

Duties for next week

Project Meeting #2
05/16 Sunday 11:59 PM

Project Milestone Due
05/16 Sunday 11:59 PM

For next Thursday:

Completed modules:
 ・ C5M1: Recurrent Neural Networks <u>(slides)</u>

Quizzes (due at 8 30am PST):
 ・ Recurrent Neural Networks
Programming Assignments (due at 8 30am PST):
 ・ Building a Recurrent Neural Network - Step by Step
 ・ Dinosaur Land -- Character-level Language Modeling
 ・ Jazz improvisation with LSTM