ECS659U/659A Coursework – The problem

The Problem

- CIFAR-10 classification
  - Dataset is readily available online
- Classify easy image in terms of 1 out of 10 classes
- Standard task for lectures & labs
- You will build a model on the training set & evaluate it on the test set

ECS659U/659A Coursework – Your Task

- Implement a **specific** model (see later) to solve the problem
  - ➢ **If you solve it using your own model you will get no marks**
- Implement the training pipeline to train the model
- Explore techniques from weeks 5-8 and from external sources
- Goal to get the highest possible accuracy

Specific note

- This is an individual assignment
- No collaboration is allowed.
- Do not use public slack channels to ask a question.
- Contact us in private

ECS659U/659A Coursework – Deliverables

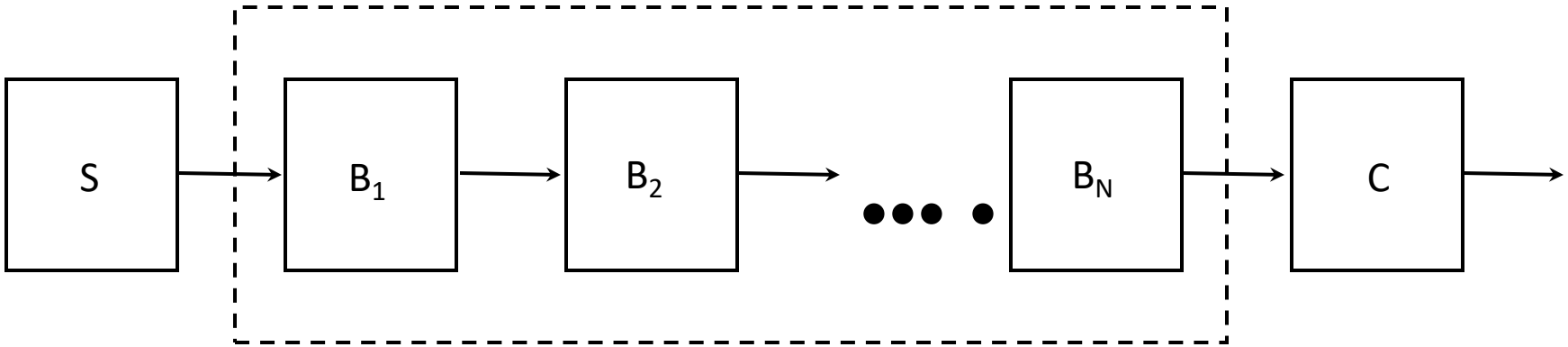
Deliverables

- They are detailed in the CW sheet.

ECS659U/659A Coursework – The Model

The Model

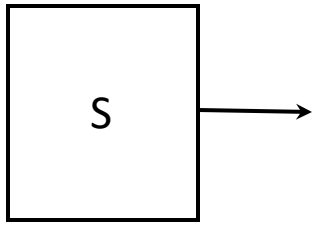
- An architecture to process images based on Convolutional Neural Networks
- Model architecture consists of Stem, Backbone (B_1, …, B_N) and Classifier.



ECS659U/659A Coursework – The Stem

The Stem

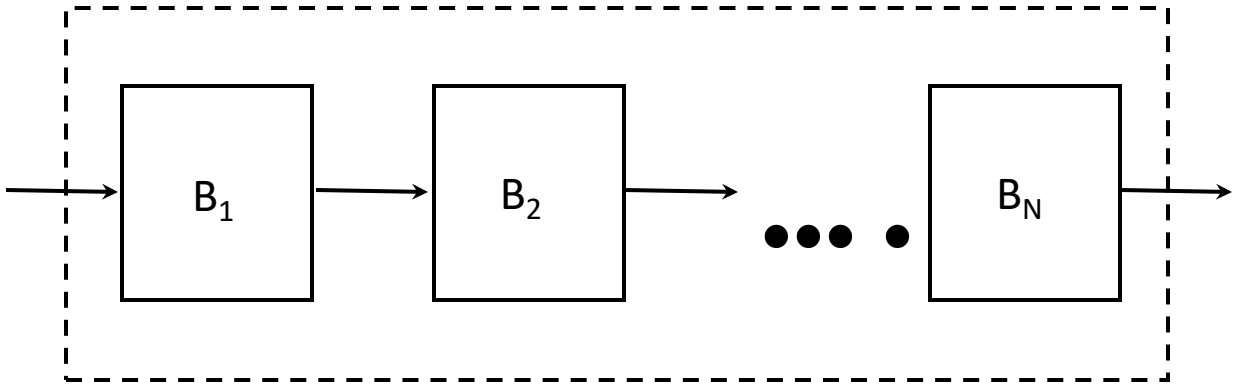
- Takes images as inputs
- Extracts a feature representation from them
  - Can simply be a Convolutional Layer



ECS659U/659A Coursework – The Backbone

The Backbone

- Consists of N Blocks. The basic (minimum) implementation for each block consists of:
- An **expert branch** predicting a vector $\boldsymbol{a} = [a_1, \dots, a_K]$ with K elements from input tensor $\boldsymbol{X}$
- A branch with K Convolutional layers which are combined using a to produce a single output $O = a_1 Conv_1(x) + \dots + a_k Conv_K(x)$
- Other components can be added based on Weeks 5-8!!

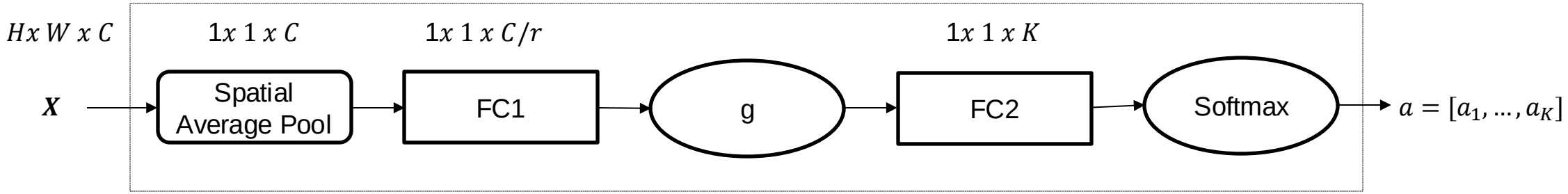


ECS659U/659A Coursework – The Block

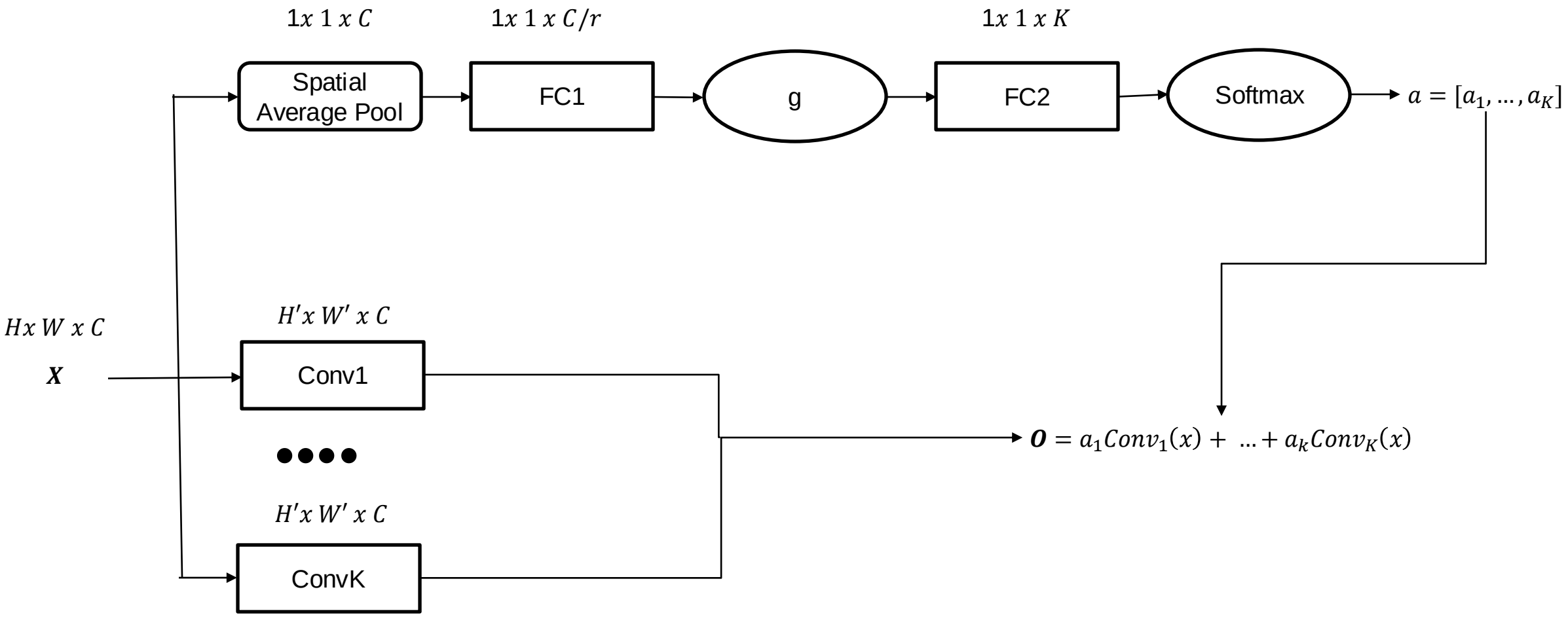
Expert branch

- The branch receives inputs from the stem for the first block and from the previous block for the rest. It generates a vector $\boldsymbol{a} = [a_1, \dots, a_K]$ with K elements from input tensor $\boldsymbol{X}$ as $\boldsymbol{a} = E(\boldsymbol{X})$**,** where E is the function that processes $\boldsymbol{X}$ for the expert branch.

- First, the input features are **spatially pooled** as $\boldsymbol{X}' = AvgPool(\boldsymbol{X})$. The pooled featured are then forwarded through **a fully connected layer (FC1)** where their channel dimensions are reduced by a factor r. Next, they are processed through a **non-linear activation function** $\boldsymbol{g}$, followed by **another fully connected layer (FC2),** which projects their channel dimensions to K. Finally, the output is forwarded through a **Softmax layer** to obtain $\boldsymbol{a} = [a_1, \dots, a_K]$.



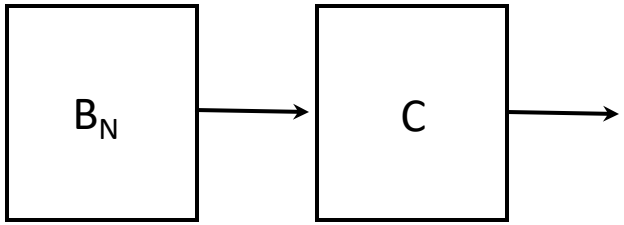
ECS659U/659A Coursework – The Block



ECS659U/659A Coursework – The Classifier

The Classifier

- Takes as input the output of the last block
- It computes a mean feature $f = \texttt{SpatialAveragePool}(O_N)$, O_N here is the output of the N_{th} block
- It passes f to a classifier
  - can be a softmax regression classifier, or an MLP
  - check also weeks 5-8



ECS659U/659A Coursework – Assessment

1. Read dataset and create data loaders: 5%
2. Create the model: 40%.
3. Create the loss and optimizer: 5%.
4. Write the training script to train the model. Provide in the report: 30%

    - the curves for the evolution of loss
    - the curves for the evolution of training and validation (test) accuracies.
    - all training details including hyper-parameters used.

5. Final model accuracy on **CIFAR-10 Validation Set**:

    - acc >90% : 20%
    - 85 < acc < 90% : 15%
    - 80 < acc< 85% : 10%
    - 70 < acc< 80% : 5%
    - acc < 70% : 0%

- Please use the coursework guidelines to create the model.
- **If you use your own architecture (or some other model) you will get no marks for Task 2, i.e -40 marks.**