

MIDDLE EAST TECHNICAL UNIVERSITY

SEMESTER I EXAMINATION 2024-2025

CENG 403 – Deep Learning - CNN Design & Transfer Learning

January 2025

TIME ALLOWED: 3 HOURS

INSTRUCTIONS TO CANDIDATES

1. This examination paper contains **SIX (6)** questions and comprises **EIGHT (8)** printed pages.
2. Answer all questions. The marks for each question are indicated at the beginning of each question.
3. Answer each question beginning on a **FRESH** page of the answer book.
4. This **IS NOT an OPEN BOOK** exam.
5. Show clear reasoning for your answers, especially intuitive explanations.
6. For architectural diagrams, draw clear components and explain design choices.
7. Connect concepts to examples discussed in lectures where relevant.
8. Explain the practical implications of design decisions.

Question 1. Position-Sensitive Convolution and Pooling (25 marks)

Based on the professor's explanation: "Vanilla convolution a CNN with vanilla convolution does not perform really well on such tasks" for position estimation.

- (a) The professor introduced position-sensitive convolution for problems where "we want to estimate a single quantity that represents the position of the object." Explain why vanilla CNNs struggle with position estimation tasks and how adding "I coordinate and J coordinate for each pixel as additional channels" solves this problem. (8 marks)

- (b) Compare pooling operations as described by the professor. Explain his statement that "complex cells do something similar to what we call pooling" in relation to neocognitron, and why "taking the maximum of the values in the receptive field works really well for many problems." (10 marks)

- (c) Analyze pooling's impact on translation invariance. The professor explained that "if the input slightly moved to the right...the maximum value is still within the receptive field." Calculate the translation tolerance for a max pooling layer with receptive field 3×3 and stride 2. (7 marks)

Question 2. Global Average Pooling and Fully Connected Layers

(22 marks)

The professor emphasized that "fully connected layers introduce some issues" and presented global average pooling as a solution.

- (a) Explain the two main problems with fully connected layers at the end of CNNs as discussed by the professor: the parameter explosion issue and the fixed input size limitation. Why does global average pooling solve both problems? (10 marks)

- (b) The professor described how global average pooling creates specialization: "different channels correspond to different objects and...we have the confidence maps highlighting the shape of the object." Explain this specialization mechanism and why "one fully connected layer is sufficient" after global average pooling. (8 marks)

- (c) The professor mentioned that "for this to work properly...we need like a similar number of channels as number of objects." Analyze the implications when this condition is not met and how the network can still "entangle" multiple objects through the policy layer. (4 marks)

Question 3. Fully Convolutional Networks

(20 marks)

Based on the professor's explanation: "Fully connected layers can be converted to convolution" for handling variable input sizes.

- (a) Describe the professor's method for converting fully connected layers to convolutional layers. Explain how "one neuron with full connectivity to its input layer" can be viewed as "a filter" and why this enables processing of higher resolution inputs. (8 marks)

- (b) The professor showed that converting FC layers to convolution produces "prediction maps" instead of single class probabilities. For a network trained on 224×224 images with 10 classes, describe what happens when you input a 448×448 image after FC-to-conv conversion. (8 marks)

- (c) According to the professor, "segmentation is a good example problem for this." Explain how fully convolutional networks enable pixel-level predictions by "sliding the whole network over the input." (4 marks)

Question 4. CNN Architecture Design Principles (25 marks)

The professor provided a "simple blueprint" for CNN design and discussed experimental findings on architecture choices.

- (a) Reproduce the professor's CNN architecture blueprint showing the sequence: "convolution convolution is followed by some nonlinearity...this can be repeated...pooling...fully connected layers." Explain why this template is widely used. (8 marks)

Draw CNN Architecture Blueprint

Show: Conv+ReLU, Pooling, FC layers

- (b) The professor discussed experimental findings: "deeper networks with smaller filters provided better results" and "depth is really critical." Explain why small filter size + deep network outperforms large filter size + shallow network, even when "one neuron in the top layer can cover the whole input range." (10 marks)

- (c) Analyze the professor's memory management strategy: "we will try to reduce dimensionality very quickly in the earlier parts of the network...information is redundant in the earlier layers." Explain why this approach works and give an example using AlexNet's "stride of four which reduces dimensionality by a factor of four." (7 marks)

Question 5. Transfer Learning Strategies

(28 marks)

The professor explained: "We are not going to design something from scratch because it is very time consuming...we will take an existing CNN and adapt that CNN to our problem."

- (a) Describe the professor's four transfer learning scenarios based on problem similarity and data size. For each scenario, specify whether to freeze/fine-tune parameters and explain the reasoning: (16 marks)

- Similar problems + small data
- Similar problems + large data
- Different problems + small data
- Different problems + large data

- (b) The professor emphasized that "earlier parts are very generic problem independent" while "later parts are problem dependent they are specific to the problem." Explain this concept using the visualization evidence he referenced and how it guides transfer learning decisions. (8 marks)

- (c) According to the professor, when fine-tuning "you need to be careful with your learning rate...you can easily disrupt the learned rates." Explain

why learning rate is critical in transfer learning and what happens if it's set too high. (4 marks)

Question 6. CNN Visualization and Interpretation (30 marks)

The professor introduced multiple visualization techniques: "Understanding predictions why it makes those predictions is important."

- (a) Describe three visualization methods explained by the professor: activation visualization, weight visualization, and finding "images that maximize a neuron." For each method, explain what patterns indicate proper network functioning vs. problems. (12 marks)

- (b) The professor explained occlusion-based analysis: "We can slide an occlusion window over the whole input...this would indicate us whether this window is important for predicting that class or not." Explain how this method detects memorization vs. proper learning. (8 marks)

- (c) Implement the professor's gradient-based saliency map approach: "If we were to assume that we can approximate the whole CNN as if it was a linear model and we take the gradient...with respect to the input." Explain the mathematical foundation and why this "highlights which parts of the input are important for a problem." (10 marks)

END OF PAPER