MIDDLE EAST TECHNICAL UNIVERSITY

SEMESTER I EXAMINATION 2024-2025

CENG 403 – Deep Learning - CNN Visualization & Classic Architectures (University Sources)

January 2025 TIME ALLOWED: 3 HOURS

INSTRUCTIONS TO CANDIDATES

1. This examination paper contains **SEVEN (7)** questions and comprises **TEN (10)** 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 all mathematical derivations clearly with proper notation.
6. For architectural diagrams, draw clear and labeled components.
7. Calculate all requested parameters and show intermediate steps.
8. Explain computational complexity where requested.

Question 1. CNN Visualization Techniques and CAM Variants (25 marks)
Based on Stanford CS231n and university computer vision course materials.

(a) Implement Class Activation Mapping (CAM) for a CNN with Global Average Pooling. Given a feature map F_k of size $H \times W$ for the k-th channel and weight w_k connecting to class c: (10 marks)

- Derive the mathematical formulation for CAM: $M_c(x, y) = \sum_k w_k \cdot F_k(x, y)$
- Explain why CAM requires Global Average Pooling (GAP) layer
- Calculate the computational complexity for generating CAM for a 512×512 image with 512 feature maps

(b) Compare CAM with Grad-CAM and Grad-CAM++. Explain the following improvements: (10 marks)

- How Grad-CAM generalizes CAM to any CNN architecture without GAP
- Why Grad-CAM++ provides better localization for objects with low spatial footprint
- Mathematical differences in weight calculation between the three methods

(c) Design an evaluation protocol for visualization methods. Propose metrics for: (5 marks)

- Quantitative evaluation of localization accuracy
- Qualitative assessment of explanation quality
- Computational efficiency comparison

Question 2. AlexNet Architecture Analysis (22 marks)
Based on D2L.ai and university deep learning course materials covering computational analysis.

(a) Analyze AlexNet's computational requirements. Given the architecture specifications: (12 marks)

- Input: 224×224×3 images
- Conv1: 96 filters, 11×11, stride 4, pad 0
- Conv2: 256 filters, 5×5, stride 1, pad 2
- FC6: 4096 neurons, FC7: 4096 neurons, FC8: 1000 neurons

Calculate:

- Memory footprint for each convolutional layer
- Number of parameters in fully connected layers vs. convolutional layers
- Which component dominates memory usage and why

(b) Evaluate AlexNet's key innovations and their impact on deep learning: (10 marks)

- ReLU activation functions: advantages over sigmoid/tanh for training speed
- Dropout regularization: mathematical formulation and overfitting prevention
- GPU utilization: architectural modifications needed for parallel processing

- Performance improvement: quantify the error rate reduction from 26.2% to 15.3%

Question 3. GoogleNet/Inception Architecture Design (28 marks)
Based on research papers and university course materials on efficient CNN architectures.

(a) Design the Inception module addressing the multi-scale processing challenge. For an input with C channels: (15 marks)

- Explain why "information can exist at multiple scales" requires different filter sizes
- Draw the naive inception module with 1×1, 3×3, 5×5 convolutions and 3×3 max pooling
- Calculate output channels: if C + C + C + C = output channels, show the growth problem
- Design the improved inception module with 1×1 convolutions for dimensionality reduction

Draw naive vs. improved inception module

Show dimensionality bottleneck solution with 1×1 convolutions

(b) Analyze GoogleNet's efficiency achievements compared to AlexNet: (8 marks)

- Parameter reduction: from 60 million (AlexNet) to 4 million (GoogleNet)
- Role of Global Average Pooling in reducing parameters
- Computational cost comparison and architectural depth (22 layers)

(c) Evaluate auxiliary classifiers in GoogleNet: (5 marks)

- Mathematical formulation of multi-loss training
- Benefits for gradient flow in deep networks
- Trade-offs in model generalization

Question 4. ResNet and Residual Learning Theory (30 marks)
Based on Microsoft Research ResNet paper and university deep learning theory courses.

(a) Analyze the degradation problem in deep networks that ResNet addresses: (10 marks)

- Why do 56-layer CNNs perform worse than 20-layer CNNs on both training and test sets?
- Mathematical explanation of gradient vanishing in very deep networks
- Distinction between degradation problem and overfitting

(b) Derive the mathematical foundation of residual learning: (12 marks)

- Given target mapping $H(x)$, show why learning $F(x) = H(x) - x$ is easier than learning $H(x)$
- Residual block formulation: $y = F(x, \{W_i\}) + x$
- Gradient flow analysis: $\frac{\partial loss}{\partial x} = \frac{\partial loss}{\partial y}(1 + \frac{\partial F}{\partial x})$
- Explain why the "+1" term prevents gradient vanishing

(c) Compare ResNet variants and their performance characteristics: (8 marks)

- ResNet-18, ResNet-34: basic blocks vs. deeper architectures
- ResNet-50, ResNet-101, ResNet-152: bottleneck blocks for efficiency
- Performance scaling: how accuracy improves with depth up to 152 layers
- Computational complexity comparison with VGG architectures

Question 5. Adversarial Attacks and Defenses (25 marks)
Based on cybersecurity research and university machine learning security courses.

(a) Formulate adversarial attack optimization problems for image classification: (10 marks)

- Targeted attack: $\min_\delta \||\delta|\|_p$ subject to $f(x+\delta) = t$ and $\||\delta|\|_\infty \leq \epsilon$
- Untargeted attack: $\min_\delta \||\delta|\|_p$ subject to $f(x+\delta) \neq y$ and $\||\delta|\|_\infty \leq \epsilon$
- Explain the role of L_p norms in constraint formulation

(b) Implement the Fast Gradient Sign Method (FGSM): (10 marks)

- Derive FGSM formula: $x' = x + \epsilon \cdot \text{sign}(\nabla_x J(\theta, x, y))$
- Explain why FGSM is effective against linear behavior in neural networks
- Calculate perturbation for a binary classification problem with cross-entropy loss
- Discuss computational efficiency compared to iterative methods

(c) Analyze defense strategies against adversarial attacks: (5 marks)

- Adversarial training: training with adversarial examples

- Defensive distillation: temperature-based softmax smoothing
- Detection methods: statistical analysis of input distributions
- Trade-offs between robustness and accuracy

Question 6. Architecture Comparison and Evolution (20 marks)
Based on comprehensive analysis from multiple university computer vision courses.

(a) Create a comparative analysis table for CNN architectures: (12 marks)

Architecture	Depth	Parameters	Top-5 Error	Key Innovation
AlexNet (2012)	8	60M	15.3%	?
VGG-16 (2014)	16	138M	?	?
GoogleNet (2014)	22	4M	?	?
ResNet-152 (2015)	152	?	?	?

Complete the table and analyze:

- Parameter efficiency trends over time
- Relationship between depth and performance
- Trade-offs between accuracy and computational cost

(b) Evaluate architectural design principles: (8 marks)

- Why smaller filter sizes (3×3) became preferred over larger ones (11×11, 7×7)
- Role of skip connections in enabling very deep networks
- Impact of global average pooling vs. fully connected layers
- Evolution from hand-crafted to learnable architectures

Question 7. Feature Visualization and Interpretability (20 marks)
Based on interpretable AI research and university courses on explainable deep learning.

(a) Design feature inversion techniques for understanding CNN representations: (10 marks)

- Formulate optimization: $\min_x \||f(x) - f_0\||^2 + \lambda R(x)$
- Explain different regularization terms $R(x)$: total variation, L_2 norm
- Implement gradient-based optimization for feature reconstruction
- Analyze why early layers preserve more spatial information than later layers

(b) Evaluate saliency map generation methods: (10 marks)

- Vanilla gradients: $S_i = \left|\frac{\partial f_c}{\partial x_i}\right|$
- Integrated gradients: addressing gradient saturation problems
- LIME (Local Interpretable Model-agnostic Explanations): local approximation approach
- Quantitative evaluation metrics for explanation quality

END OF PAPER