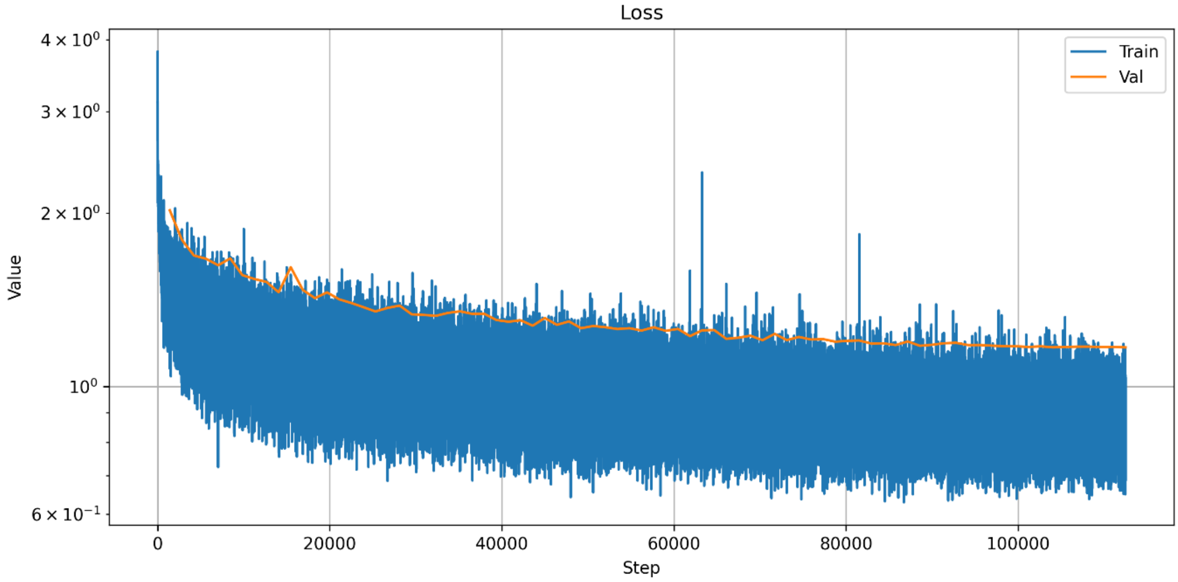
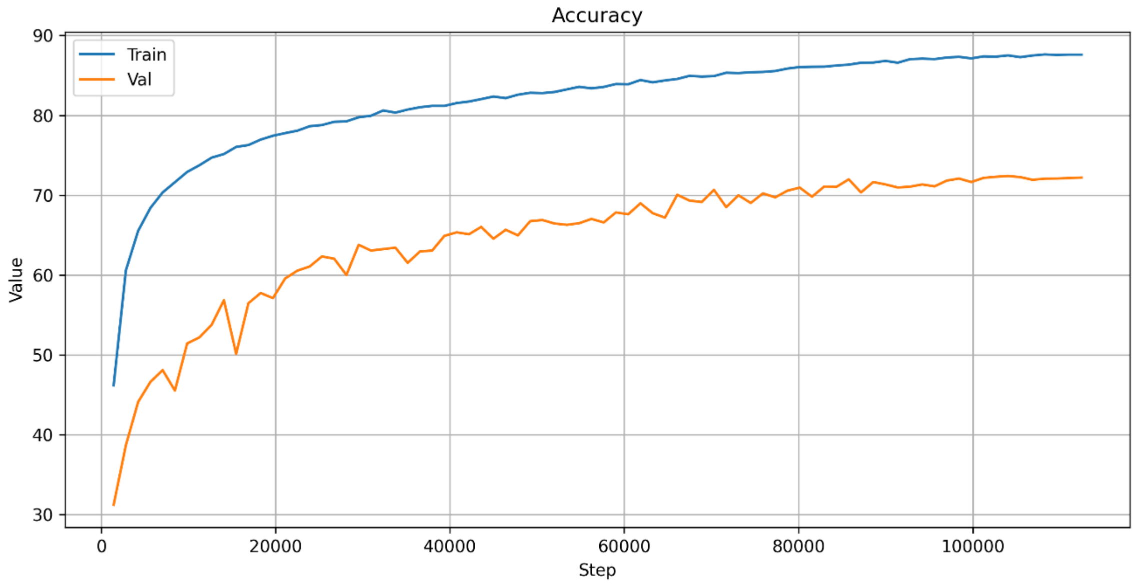
Report

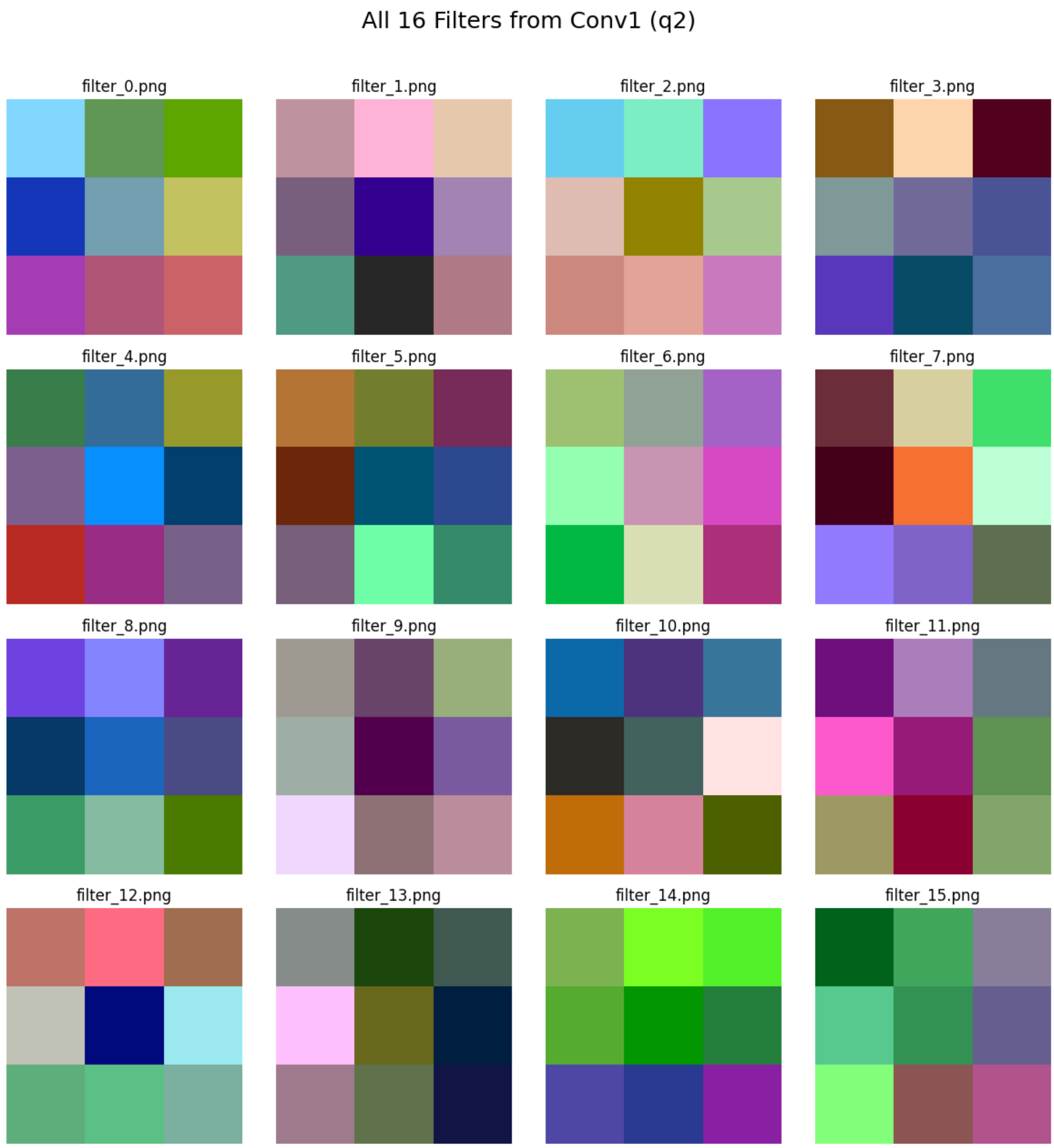
1) **Report the validation accuracy in your report**

The validation accuracy achieved is ~72.5 %





2. Visualize all the filters from the first layer of your best model. Do these filters remind you of any patterns related to any of these CIFAR-10 classes?



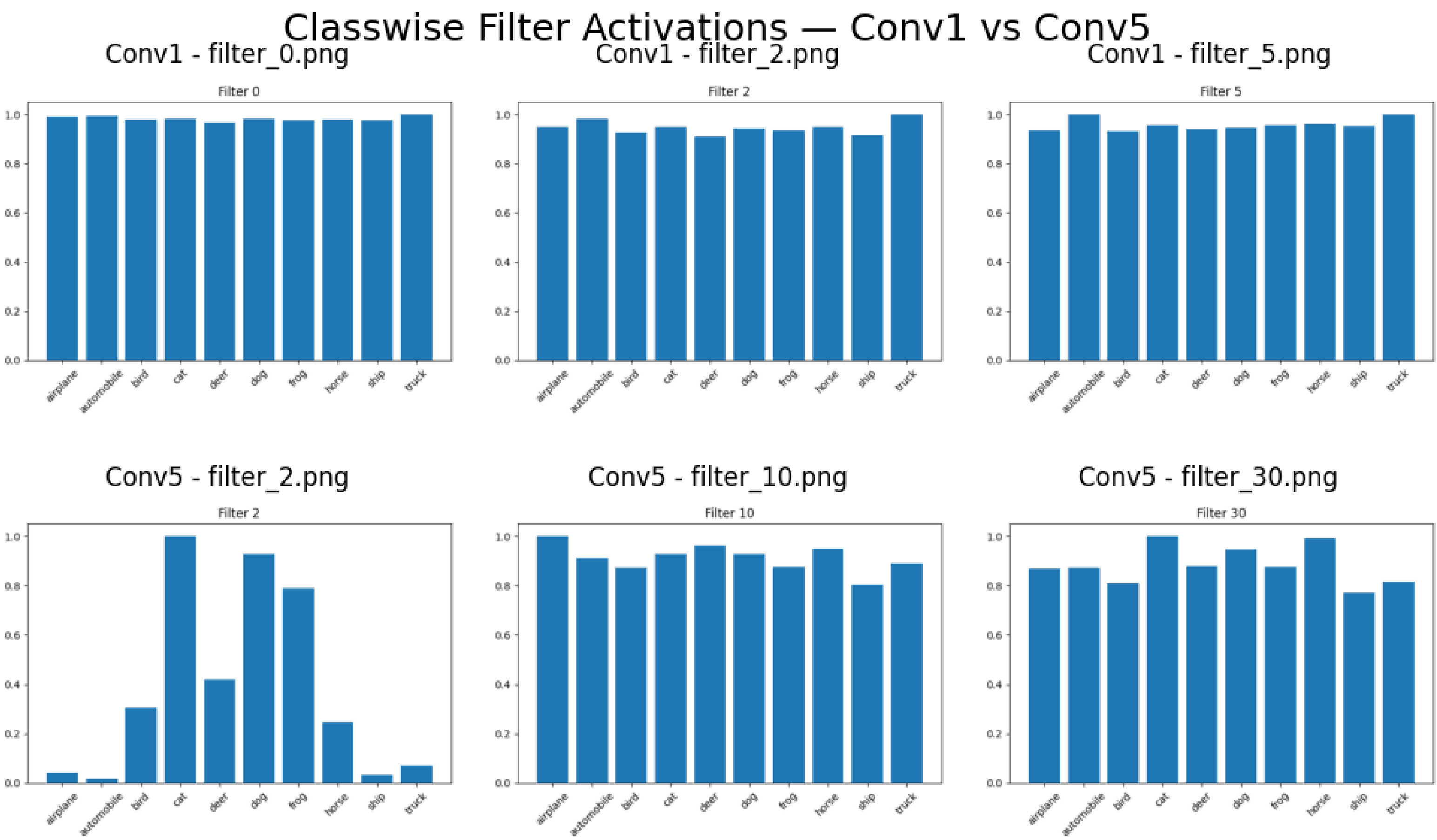
- Some filters appear to highlight sharp transitions between light and dark — common in edge detection.
- These help the model detect boundaries of objects like airplanes, automobiles, or trucks.
- A few filters seem dominantly red or green.
- Such filters might respond strongly to objects like:
  - Frogs or birds (green backgrounds)
  - Deer or dogs (brown tones)
  - Ships (blue/gray tone)
- Some filters show patterns that look like corner detectors or textures.

- These might be useful for differentiating cats vs. dogs, or automobiles vs. trucks based on texture or contour.

Yes , these filters relate to CIFAR-10 classes — but indirectly.

- These filters don't look like planes or ships — but they capture the building blocks of those images.
- For example:
    - To recognize a ship, the network might combine edge + color filters to find hull outlines on water.
    - For cats, texture and color filters help identify fur patterns and shapes.

3. Include a few bar plots corresponding to filters from the first and the final convolutional layers.



4. Explain what difference you note between the bar plots from these two layers. How can you use them to explain what your model has learned?

Filters from Conv1 (First Layer):

- The bar plots are **fairly uniform** across all 10 classes.
- This suggests that these filters are **general-purpose**, detecting:
    - Edges, textures, color gradients
    - Common features found across multiple object types
- They are **not specialized** — expected behavior for early layers.

Filters from Conv5 (Last Layer):

- The bars are **highly class-specific**:
    - Some filters respond strongly to one or two classes (e.g., cat or dog).
    - Others are almost inactive for unrelated classes.
- This shows that later filters have learned to **specialize** in identifying patterns **unique to certain classes**.

These bar plots visually demonstrate the **hierarchical learning** structure of CNNs. Filters from early layers (e.g., Conv1) are activated similarly across all classes, indicating that they capture general features such as edges and colors. In contrast, filters from deeper layers (e.g., Conv5) show class-specific activation patterns, implying that the model has learned to associate certain filters with distinct object types. This hierarchical learning is a key reason why CNNs are effective at classification tasks.

5. Analyze the plots corresponding to the filters from the first and the last layers. Include a descriptive analysis in your report. Hint: You can look into the variance across the classes for each filter. Are they uniform? What do activations that are uniform across the classes mean?

When a filter in a convolutional neural network (CNN) produces similar activation levels for all classes, we say that its activations are uniform across the classes.

Conv1 (First Layer) — General Feature Extractors

- Observation: The bar plots from Conv1 filters exhibit **low variance** across the 10 classes.
    - For example, filters such as filter_0, filter_2, and filter_5 show bars of nearly equal height.
- Interpretation: These filters are not class-specific.
    - They detect basic image patterns like:
        - Horizontal and vertical edges
        - Color gradients or textures
        - Local contrast features
    - Such patterns are common across all images in CIFAR-10, regardless of class.

- Conclusion: **Uniform activation** across classes indicates that Conv1 filters function as shared building blocks used to construct higher-level features downstream.

Conv5 (Last Layer) — Class-Specific Feature Detectors

- **Observation**: The bar plots from Conv5 filters demonstrate **high variance across classes**.
    - Some filters (e.g., filter_30, filter_50) show strong activation for only one or two classes.
    - Other filters remain nearly inactive for unrelated categories.
- **Interpretation**: These filters have become **specialized** to activate in the presence of high level class-specific cues.
    - For example:
        - A filter may fire strongly on "ship" images due to detecting waterline features.
        - Another may activate on "cat" due to detecting fur texture or face shape.
- **Conclusion**: High classwise variance indicates that the filter plays a **discriminative role** — helping the classifier distinguish between classes.