Project: Anomaly Detection and Severity Classification of Diabetic Retinopathy Using Deep Learning

Dawood AL CHANTI
Enseignant-Chercheur
Grenoble INP - Phelma, GIPSA-Lab
dawood.al-chanti@grenoble-inp.fr
(5PMBMLD0) BIOMED 3A, S9

1 Introduction

Diabetic retinopathy (DR) is a complication of diabetes that can lead to vision loss. Early detection of DR from retinal fundus images is crucial for preventing blindness. In this project, you will use the APTOS 2019 Blindness Detection dataset (Kaggle) to develop deep learning models that perform two tasks: (1) detect anomaly (binary classification: DR vs. no DR) and (2) classify DR severity (ordinal multi-class from 0 to 4). This combines anomaly screening with fine-grained grading. The motivation is to apply state-of-the-art convolutional neural networks (CNNs) to a medical imaging task and to emphasize model interpretability over raw performance, reflecting the needs of clinical decision support.

2 Dataset and Task

The APTOS 2019 Blindness Detection dataset contains 3662 retinal fundus images labeled with DR severity (0 = No DR, 1 = Mild, 2 = Moderate, 3 = Severe, 4 = Proliferative). It is accessible from Kaggle https://www.kaggle.com/c/aptos2019-blindness-detection. You should split the data into training, validation, and test sets (e.g. 70/10/20% or use cross-validation) and analyze class balance. Provide visualizations of the class distribution and example images from each DR grade.

3 Methods

In this project, you will rely on the idea of transfer learning, were you will use pretrained CNNs (e.g. ResNet50, EfficientNetB3/V2, DenseNet121, etc.)

via Keras/TensorFlow or PyTorch. Follow a transfer learning workflow: load a model pretrained on ImageNet, or any model of your choice, replace the classification head for 2 or 5 outputs, and fine-tune the network on APTOS. Suggested steps include:

- Preprocess images (e.g. resize to 224×224 or model-specific input size, normalize pixel values to [0,1] or ImageNet means) and apply optional augmentation (random flips, rotations, brightness shifts) to increase robustness.
- Initially, freeze convolutional layers and train only the new head; then unfreeze and fine-tune some deeper layers.
- Use callback techniques such as early stopping and learning-rate reduction to prevent overfitting.
- Ensure reproducibility (set random seeds, document data split).

For reference, recent work used multiple pretrained CNNs (EfficientNetB3, EfficientNetV2B1, RegNetX, etc.) on this dataset. You may also explore architectures like DenseNet or MobileNet if desired, and/or compare to ResNet/EfficientNet archiectures.

4 Evaluation Metrics

Evaluate your models using standard classification metrics. For the binary DR detection, report accuracy, precision, recall, F1-score, and plot the confusion matrix. For the 5-class severity classification, also report accuracy, per-class precision/recall, F1-scores, and a 5×5 confusion matrix. In addition, compute the *Quadratic Weighted Kappa* (QWK), which is a measure of agreement sensitive to the ordinal severity ranking. QWK is commonly used in DR grading challenges (e.g. 0.93 was achieved in one study). Explain any class imbalance handling (e.g. class weighting or sampling) and justify metric choices.

You should visualize:

- data distributions (histogram of grades)
- a few sample images per class
- training and validation loss/accuracy curves for each model,
- confusion matrices (either heatmaps or tables).

These figures help diagnose issues like under/over-fitting and class confusions.

5 Bonus: Model Interpretability

Beyond numerical performance, emphasize interpretability. Generate visual explanations (e.g. Grad-CAM or Grad-CAM++, or any saliency maps of your choice) for representative test images to highlight which retinal regions influence the model's predictions. Comment on whether the model is focusing on clinically relevant features (such as lesions or hemorrhages). Describe the interpretation method and discuss its results in your report. Interpretable AI is critical in medical imaging to build clinician trust [2], so allocate significant analysis here.

6 Deliverables

You must submit a concise technical report (maximum 6 pages, not including references). The report should include:

- **Abstract:** Brief overview of problem, methods, and key results.
- **Introduction:** Context of diabetic retinopathy, task objectives, and motivation.
- **Methods:** Description of data preprocessing, model architectures, training procedure (transfer learning, fine-tuning details).
- **Results:** Quantitative evaluation (metrics, tables/figures of performance), visualizations (learning curves, confusion matrix, etc.).
- **Discussion:** Interpret the results, including interpretability analysis; discuss limitations and potential improvements (see next section).
- **Conclusion:** Summary of findings and takeaways.
- **References:** Cite all sources in BibTeX format.

Include all requested figures (example images, plots) and clearly label sections. Use captions for tables/figures and mention them in text.

7 Literature Review

Include a brief literature review (1–2 paragraphs) citing 2–3 recent relevant papers in DR classification. For example, Alyoubi *et al.* (2021) used CNNs (called CNN512) and a YOLOv3 model to classify DR into five stages on APTOS, reporting around 84% accuracy [1]. Youldash *et al.* (2024) applied multiple pretrained CNNs (EfficientNet variants and RegNet) to APTOS for binary and multi-class DR detection. Chetoui and Akhloufi (2020) developed an end-to-end DR detection model with an explanation module (via Grad-CAM) to highlight salient DR features [2]. Summarize each reference's approach and how it relates to your project. Provide the full BibTeX entries for these citations in the references section.

8 Discussion, Limitations, and Future Work

In your report discussion, critically analyze your approach. Note any limitations (e.g. small dataset size, class imbalance, image quality variability). Suggest alternative strategies not pursued and explain why they were omitted. Propose future improvements (e.g. collecting more data, advanced augmentation, multi-task learning, federated learning, or additional interpretability techniques). This demonstrates understanding of the broader context and remaining challenges.

Dataset and Challenge Link

The APTOS 2019 Blindness Detection dataset and challenge details are available on Kaggle: https://www.kaggle.com/c/aptos2019-blindness-detection. You are expected to download the data from this source.

References

[1] Alyoubi, W. L., Abulkhair, M. F., and Shalash, W. M., Diabetic Retinopathy Fundus Image Classification and Lesions Localization System Using Deep Learning, *Sensors*, vol. 21, no. 11, pp. 3704, 2021, doi 10.3390/s21113704.

[2] Chetoui, M., and Akhloufi, M. A., Explainable end to end deep learning for diabetic retinopathy detection across multiple datasets, *Journal of Medical Imaging*, vol. 7, no. 4, article 044503, 2020, doi 10.1117/1.JMI.7.4.044503.

[3] Alshammari, G., and Alqahtani, M., Early Detection and Classification of Diabetic Retinopathy A Deep Learning Approach, *AI*, vol. 5, no. 4, pp. 2586 to 2617, 2024, doi 10.3390/ai5040125.