Information and Networking Event Horizon Europe 2023-2024 Calls Co-Funded by the Government of India (DST)

HORIZON-CL4-2024-HUMAN-03-02: Explainable and Robust AI

24 May 2024

- TITLE of talk (edit): Explainable AI Model for Geographic Atrophy Detection

- Name of presenter Dr Pushan Kumar Dutta
- Email – pkdutta@kol.amity.edu/ pkdgeoindia@gmail.com
- Name of Organisation Amity University Kolkata (Academic Institution)
- Country : India
- Your contact details: 9836331735
- Your web url: https://orcid.org/0000-0002-4765-3864
- Linkedin: https://www.linkedin.com/in/pushanduttaeng/

Topic to be addressed

- This study demonstrates the potential of explainable AI (XAI) for improving the clinical adoption of deep learning models in ophthalmology. Future work will focus on integrating explainability with other retinal diseases and optimizing the system for teleophthalmology applications.
- XAI algorithms can analyse retinal images to identify disease- specific features and patterns, while providing a clear explanation of the diagnostic process, which not only improves diagnostic accuracy but also streamlines decision- making, enabling faster and more efficient treatment.
- Geographic Atrophy (GA): A growing eye disease causing vision loss. Early detection is critical for treatment.
- Colour fundus photography (CFP): A widely available and affordable imaging technique for detecting GA.
- Challenges with AI for GA diagnosis: Previous AI models using CFP lacked accuracy and explainability.
- This study introduces a new AI model for GA detection with CFP images

Process

- 1. Retinal imaging was performed on each patient using a fully automated confocal scanning ophthalmoscope (Eidon, CenterVue Spa) that uses a white LED light source and captures 60-degree, 14-megapixel color retinal images through a non-mydriatic pupil.
- 2. Color fundus images centered on the foveal midpoint were acquired for both eyes of each participant under their natural pupil size. A single image was captured for each eye.
- 3. A technician ensured that the acquired images were of gradable quality. Blurred or defocused images due to severe cataracts or keratitis were excluded from further analysis.
- 4. All fundus images were extracted as JPG files and anonymized to ensure privacy and confidentiality.
- 5. The anonymized fundus photographs were then uploaded to an AI system for further analysis and evaluation.
- In summary, a confocal scanning ophthalmoscope was used to capture high-resolution color fundus images of the retina for each eye of the participants. The images were quality checked, anonymized, and then prepared for analysis by an AI system.

Solution

- We propose a novel AI model for GA detection using retinal images captured by a non-mydriatic confocal scanner.
- Key Features:
- High Accuracy: The model leverages a pre-trained Deep Convolutional Neural Network (Efficientnet_b2) for accurate image classification.
- Explainability: A GradCAM module generates heatmaps highlighting image regions most influential for the model's decision, fostering trust and understanding for clinicians.
- Process:
- 1. Image Acquisition: Standardized retinal images are captured from participants using a confocal scanner.
- 2. Image Preprocessing: Images are anonymized and converted to JPG format.
- 3. Classification: The Efficientnet_b2 model classifies images into three categories: healthy, GA, or other retinal conditions.
- 4. Explainability: GradCAM generates heatmaps visualizing the image regions critical for the model's classification.
- Benefits:
- Improved Screening: Automated analysis facilitates faster and more accessible GA screening.
- Informed Decisions: Explainable AI empowers clinicians to understand the model's reasoning, leading to more confident diagnoses.
- Enhanced Treatment: Early detection enables timely intervention with new GA therapies, potentially preserving vision.

1. **Input Image:** The model takes a retinal image captured using a non-mydriatic confocal scanner as input.
2. **Feature Extractor:** The Efficientnet_b2, a pre-trained deep convolutional neural network (CNN), processes the input image to extract relevant features. Here, the final few layers are purposefully removed to prevent overfitting to the training data (retinal images) as the model was pre-trained on a broader dataset (ImageNet).
3. **Output:** The model outputs a classification label for the image, indicating one of three categories: healthy, GA, or other retinal conditions.
4. **Grad-CAM Module:** This module acts as the explainability component. It takes the feature maps generated by the final convolutional layer of the feature extractor and calculates weights based on the gradients. These weights are used to create a heatmap highlighting the image regions that significantly influenced the model's classification decision.
5. **Heatmap:** This visual output aids in understanding the model's reasoning behind its classification. By looking at the heatmap, clinicians can see which parts of the image were most critical for the model's decision (e.g., areas with potential signs of GA).

- **Enhanced Study on Perception:**
- **User Interface:** The user interface should be designed to seamlessly integrate the heatmap alongside the classified image. This allows for easy comparison between the model's output and the highlighted regions.
- **Interactive Heatmap:** Consider developing an interactive heatmap where hovering over specific areas reveals additional information about the corresponding image features.
- **Clinician Feedback:** Incorporate feedback from ophthalmologists and optometrists regarding the clarity and usefulness of the heatmaps. This feedback can guide further refinement of the explainability module.

1. Data Preprocessing:

- Perform data augmentation techniques (e.g., rotation, flipping, scaling) on the training dataset to increase diversity and improve model generalization.
- Normalize the pixel values of the color fundus photographs.

2. Model Architecture:

- Develop an ensemble learning approach by combining multiple CNN models with different architectures (e.g., EfficientNet, ResNet, DenseNet).
- Each CNN model will be trained independently on the augmented training dataset.

3. Training Process:

- Split the augmented dataset into training and validation sets.
- Train each CNN model separately using the training set and monitor performance on the validation set.
- Employ various regularization techniques (e.g., dropout, early stopping, weight decay) to prevent overfitting.
- Optimize model hyperparameters (e.g., learning rate, batch size) using techniques like grid search or Bayesian optimization.

4. Ensemble Learning:

- Combine the predictions of the trained CNN models using an ensemble strategy, such as majority voting or weighted averaging.
- Explore different ensemble techniques, such as bagging or boosting, to improve the overall performance.

5. Explainability:

- Implement an explainability technique, such as Grad-CAM or LIME (Local Interpretable Model-Agnostic Explanations), to generate heatmaps or explanations for the ensemble model's predictions.
- Visualize the regions of the fundus images that contributed most to the GA detection.

- Dr Pushan Kumar Dutta, Ph.D.
- **Core Competencies:**
- - Designing and conducting experiments
- - Analyzing complex data sets
- - Developing new theories and models
- - Authoring research papers and reports
- - Expertise in field of computer vision/medical imaging Core
- Research Tasks: - Formulating novel research questions and hypotheses
- - Planning and executing studies to test hypotheses
- - Applying analytical techniques to datasets
- - Interpreting results and drawing insights
- - Communicating findings through publications/presentations

- **Support Needed**
- Assistance with literature review and referenced information gathering
- - Support in processing and analyzing large datasets
- - Feedback on clarity and flow of research writing
- - Suggestions for enhancing study design and methodology
- - Explaining complex concepts in easy to understand manner

Keywords

- 1. Geographic Atrophy (GA)
- 2. Explainable AI (XAI)
- 3. Color Fundus Photography (CFP)
- 4. Automated Screening
- 5. Early Detection
- 6. Age-related Macular Degeneration (AMD)
- 7. Retinal Imaging
- 8. Computer Vision
- 9. Deep Learning
- 10. Interpretable Machine Learning
- 11. Clinical Decision Support
- 12. Healthcare Accessibility
- 13. Cost-effective Screening
- 14. Vision Preservation

Specialization

1.Ophthalmologists/Retina Specialists:

- - Skills: Clinical expertise in GA diagnosis, retinal imaging interpretation, and patient management.
- - Role: Provide medical guidance, validate model performance, interpret results, and facilitate clinical adoption.

2. Computer Vision/Machine Learning Experts:

- - Skills: Proficiency in deep learning, computer vision, and explainable AI techniques.
- - Role: Develop and optimize the XAI model, implement explainability methods, and ensure model robustness.

3. Software Engineers:

- - Skills: Experience in software development, user interface design, and deployment of AI systems.
- - Role: Develop user-friendly software applications, integrate the XAI model, and ensure seamless integration with existing healthcare systems.

4. Healthcare Administrators/Policymakers:

- - Skills: Understanding of healthcare systems, resource allocation, and policy implementation.
- - Role: Facilitate the adoption and integration of the XAI model into clinical workflows, ensure compliance with regulations, and promote accessibility.

5. Data Scientists/Researchers:

- - Skills: Expertise in data acquisition, processing, and analysis, as well as research methodology.
- - Role: Curate and preprocess retinal image datasets, perform data analysis, and contribute to research publications.

6. Patient Advocacy Groups:

- - Skills: Understanding of patient needs, concerns, and perspectives.
- - Role: Provide feedback on the model's usability, interpretability, and potential impact on patient care.

THANK YOU