PROJECT ABSTRACT

Malaria Detection using CNN

CNN classifier that distinguishes parasitized from uninfected red blood cells in thin-smear microscopy images, with a smear-scanner demo — delivered as a built-to-order student project (research prototype, not a medical device).

01. Title

Malaria Detection using CNN

02. Introduction / Background

Malaria still kills hundreds of thousands of people each year, and the gold-standard diagnosis — manual examination of Giemsa-stained thin blood smears under a microscope — is slow, error-prone and depends on scarce expert microscopists. Convolutional neural networks, which learn texture and shape cues directly from pixels, are a natural fit for the cell-level classification at the heart of this workflow. This project builds a complete student system: a CNN trained on the public NIH malaria cell-image dataset, with a smear-scanner style demo that classifies a field of segmented cells. Medical-prototype disclaimer: this project is an academic prototype and is not a diagnostic or clinical device. It requires oversight by qualified medical professionals and is never a substitute for clinical diagnosis.

03. Problem Statement

A microscopist must examine thousands of red blood cells per slide to find parasites — a 10-30 minute task per patient that suffers from fatigue and inter-observer variation, and expert readers are unavailable in many endemic regions. The project addresses this by building an automated cell classifier that scores each segmented cell as parasitized or uninfected and summarizes the field, acting as a triage aid whose every decision must be confirmed by a trained microscopist.

04. Objectives

- Train a convolutional neural network on the NIH malaria cell-image set (27,558 segmented cells, balanced parasitized/uninfected) for binary classification.
- Reach a design target in the high-90s validation accuracy on the balanced dataset, with the notebook logging accuracy, precision, recall, F1 and ROC-AUC.
- Build a smear-scanner demo: cell grid, per-cell prediction, field summary and activation visualization.

- Deliver the complete student kit: source code, trained weights, report, PPT and viva Q&A document.

05. Proposed Methodology

The system has three stages. (1) Data pipeline: NIH cell images are resized to 128x128 RGB, stain-normalized, and augmented with rotations, flips, zooms and stain-color jitter; splits use patient-level grouping so cells from the same patient never appear in both train and test. (2) Model: a 4-block CNN (32->64->128->256 filters with batch normalization and max-pooling) followed by global average pooling, a dense layer with dropout and a sigmoid output, trained with binary cross-entropy and the Adam optimizer. (3) Demo: the scanner view classifies each cell in a field and aggregates the count; an analysis view shows filter activations and parasite-chromatin markers on selected cells.

06. System Architecture / Working

Cell image (128x128 RGB) -> stain normalization -> convolutional blocks (conv-batchnorm-ReLU-maxpool x4) -> global average pooling -> dense(128) + dropout -> sigmoid -> P(parasitized). The scanner runs this per cell over a field and renders a color-coded grid; the analysis view overlays activation-derived markers on a chosen cell and plots mid-level feature responses (chromatin response, ring-form texture, edge density).

07. Hardware / Software Requirements

- Python 3.10, TensorFlow/Keras or PyTorch, OpenCV, NumPy, pandas
- scikit-learn (metrics, patient-level splits), Matplotlib/Seaborn
- Flask demo web app (smear scanner + cell analysis views)
- GPU recommended for training; inference runs on CPU

08. Expected Outcomes

A trained cell classifier with a design target in the high-90s validation accuracy on the balanced NIH set (measured in the buyer's own run, never claimed in advance), a working scanner demo that flags parasitized cells in a field, activation visualizations for the report, and full documentation. The system is an educational research prototype, not a diagnostic device.

09. Applications

- Teaching material for CNNs, medical-image classification and handling stain-variation in microscopy
- Prototype triage aid for thin-smear workflows (non-clinical)
- Reference build for other blood-cell classification tasks

- Demonstration of patient-level splitting to avoid data leakage

10. Future Scope

- Parasite life-stage classification (ring, trophozoite, schizont, gametocyte)
- Whole-slide scanning with cell segmentation (U-Net) instead of pre-segmented cells
- Smartphone-microscope deployment pipeline (as in the NLM Malaria Screener work)
- Multi-species classification (P. falciparum vs P. vivax)