Utilizing Deep Learning to Facilitate Diagnosis of Look-Alike Leukemia Subtypes

Chan, Tessla (School: Roslyn High School)

B-cell acute lymphoblastic leukemia (B-ALL) and acute myeloid leukemia (AML) are the 2 most common leukemias. The American Cancer Society estimates that there will be 59,610 leukemia cases and 23,710 deaths in 2023. Early detection of leukemia subtypes is critical in prescribing effective treatment strategies to maximize survival rates. Past studies using machine learning (ML) for diagnosing leukemia have achieved significant results, but very few studies were extended to cover the detection of specific leukemia subtypes. B-ALL and AML subtypes are difficult to distinguish because they have similar morphology. The objective of this study is to develop a novel dual-input model to distinguish subtypes of B-ALL and AML. Transfer learning will be utilized, and different classification algorithms: Fully-connected (FC), FC + Batch Normalization (BN), Global average pooling (GAP) will be tested on top of 3 pre-trained CNNs (VGG16, DenseNet201, InceptionResNetv2). Image segmentation will identify malignant cells, and data augmentation will apply transformations to increase the number of training samples. The model's hyperparameters will be finetuned while checkpoints and early stopping callbacks track loss and mitigate overfitting. Of all CNNs tested, DenseNet201 with the GAP classifier achieved optimal test accuracy of 96.62% with an AUC score of 0.9989. An app called LST Detector will be developed for public use and to create a community-supported repository of blood smear images for the model to undergo further training. This model is the first to successfully identify all morphologically similar subtypes of B-ALL and AML leukemias.