Introduction

Haemoglobin is the protein found in red blood cells that helps deliver oxygen to our tissues (Billett, 1990). Genesis 1 is an AI-powered app that uses a selfie image to predict one's blood haemoglobin count based on the blood flow in the lips without pricking the lips (Genesis1, n.d.).

Our objective is to build a proprietary algorithm that estimates haemoglobin reading ("HgB") from lip images and metadata with ≤ 0.8 g/dL MAE while remaining robust across devices, lighting and skin tones.

Related Works

It is shown that haemoglobin levels can be determined by our skin colors, such as from our eyelids. A paper from Chambers et. al (2021) used images of eyelids taken from iPhones to predict the levels of haemoglobin in a person. Preliminary results also showed that they only used images without the use of flash lighting in their predictive model as the error for the predicted haemoglobin levels was significant. They also showed that the hue, saturation and value (HSV) was more important than its RGB value in determining the haemoglobin levels.

Approach

We loaded all the lip images and extracted haemoglobin values from the filenames. We split the data to 80% training data and 20% testing data to ensure robust model evaluation. We decided to use Convolutional Neural Networks (CNN) to predict the haemoglobin levels based on lip colors as it is well-suited for image-based regression tasks. CNN passes an input image through a convolutional layer, or a filter which automatically learns and extracts spatial features such as edges, textures and colors, to produce a feature map. An activation function is applied on our image's feature map to obtain an activation map, where a pooling and flatten layer can be applied on to obtain inputs for a deep neural network model to learn the relationships between hue, color and haemoglobin levels.

ResNet-18
ResNet-18 is the smallest and most lightweight model among the ResNet family. It is a popular model for image classification that balances speed, accuracy and complexity. ResNet uses skip connections that bypass layers to allow the gradients to flow more smoothly, even for models with many layers. With only 18 layers, its simplicity is suitable for our considerably small dataset to prevent overfitting of our model. It has also been proven to perform well on small datasets (Roboflow, 2025).

Our model has a size of 42.7MB and a processing speed of around 35ms. These results make it suitable for the model to be run on a smartphone that has limitations in computational resources and storage, allowing for users to get fast results from their phones.

Preprocessing

Best Model

In order to find the model with the best accuracy, we ran the data through 100 epochs using the Adam optimizer and L1 loss (Mean Absolute Error). The learning rate was reduced whenever the performance of the models started to stagnate. After the model training, we managed to get the the model with the lowest Validation MAE score/

Our evaluation metrics showed that it is able to predict haemoglobin levels from unseen images accurately and reliably. Our error metrics show that on average, our predicted haemoglobin values are within 0.8g/dl of the actual values. The high R-squared value shows that our model can explain a significant portion of the variance in haemoglobin levels. These results show that we can provide users with fast and accurate haemoglobin results just by providing images of their lips from their smartphone.

Feature Extraction

The architecture modifies the standard ResNet18:

- Removing the classification head (originally 512→1000 for ImageNet classes)
- Replacing with a regression head (512→1 linear layer)
- Applying squeeze() to output scalar hemoglobin values

Optimisation Strategy

The training configuration demonstrates sound practices:

- Adam optimizer (lr=1e-4)
- Validation-based checkpointing

Input Processing

The current implementation processes

- Resize operations to 224x224
- normalization (mean=[0.485, 0.456, 0.406], std=[0.229, 0.224, 0.225])

Data Augmentation Strategy

These augmentations help the model generalise to better by creating variation, lighting changes, and other environmental factors given a data set of 32 images.

- Random horizontal/vertical flips - realistic for lip imaging variations
- Color jitter - to simulate lighting variations
- Random resized crop - handles different capture distances

- Random rotations - accounts for device orientation differences

References

Mahmud, S., Donmez, T. B., Mansour, M., Kutlu, M., et al. (2023).
Anemia detection through non-invasive analysis of lip mucosa images.
Frontiers in Big Data, 6, Article 1241899. https://doi.org/10.3389/fdata.2023.1241899

Roboflow. (2023, August 2). *ResNet-18: Architecture and applications*. Roboflow Blog.
https://blog.roboflow.com/resnet-18/

Suner, S., Rayner, J., Ozturan, I. U., Hogan, G., Meehan, C. P., Chambers, A. B., et al. (2021).
Prediction of anemia and estimation of hemoglobin concentration using a smartphone camera.
PLoS ONE, 16(7), e0253495. https://doi.org/10.1371/journal.pone.0253495