

ABOUT UHAI

Universal Health Artificial Intelligence (UHAI) is a decentralized platform that processes, stores and unifies electronic medical information on the blockchain where AI is used to detect anomalies and predict health outcomes. UHAI is powered by UHAI cryptographic utility tokens that enable participants to earn tokens in exchange for participation. Security and privacy are first priorities.

Please visit us at www.uhai.io

FutureHack Hackathon

The UHAI Breast Cancer Detection Project is an end-to-end offering that can determine if a mass lesion on an x-ray image is benign or cancerous. To achieve this analysis, we leveraged IPFS (a decentralized file storage) and Ethereum blockchains in combination with our internally tuned machine learning algorithms. Currently users can receive almost as good as a radiologist diagnosis by simply uploading an mammography (68% accuracy developed during the Hackathon) which can be improved above 90% based on recent studies³. The results are returned within seconds.



IMPACT: We are providing solutions under SDG 3, Health and Wellbeing. We have chosen breast cancer because it is the number one killer of women globally at 800k deaths annual. As well, it's the number one new cancer case in the world, at 30%. Breast cancer affects every mother, daughter, aunt— every woman on the planet. For the very first time, patients can upload her medical information to a platform that will give instant diagnosis. The average error rate among radiologists is around 30%³. With the help of large datasets and deep learning algorithms our system can help doctors and radiologists to detect malignant breast lesions and decrease this ratio.

BLOCKCHAIN: We use the IPFS (decentralized file storage) to store image files and the Ethereum blockchains to store transactional information and metadata.

SCALE: The system will scale as fast as the blockchain network. Plus, the front-end is a webservice and can be powered up by renting computation from provider. The platform can be globally scaled instantly.

DESIGN: The system has many components, designed to work separately or collaboratively. We design the system to blockchain agnostic, which effectively shields us from performance issues of certain blockchains. If a blockchain cannot scale as fast as we need, we can move to a better one.

WALKTHROUGH:

- 1) User user logs into web
- 2) User uploads x-ray image
- 3) Data is sent to IPS for storage. System retrieves hash.
- 4) AI engine runs analysis on x-ray.
- 5) Results and history are sent back to the user.
- 6) Results + IPFS hash are written onto the Ethereum blockchain.

OVERALL ARCHITECTURE DIAGRAM

Please see technical document.

AI Engine Logic Description

Deep Learning - Convolutional Neural Network_7_final.ipynb

Inception V3 Deep Neural Network Used to Predict Breast Lesion Diagnosis

Our model uses the Inception V3 Convolutional Neural Network pretrained on Imagenet, retraining the top 3 Inception layers to fit the Digital Database of Screening Mammography¹ (DDSM) images segmented via radiologist notation and cropped into a Region of Interest (ROI) by the Cancer Imaging Archive (TCIA) Public Access².

The Inception V3 Neural Network was designed by Google and initially trained on the ImageNet Large Visual Recognition Challenge. It consists of multiple convolutional, pooling, and concatenation batch normalization nodes, in addition to 11 inception models that contain parallel pathways of convolution and pooling nodes.³

Training and model evaluation steps:

Loaded data and kernel on AWS Ubuntu GPU Instance in order to reduce large training times.

As detailed in the jupyter notebook, we first converted the images into numpy array format, converted grayscale to RGB format, resized images to 299x299, and normalized values to 0-1, and coded response binary data ("BENIGN" or "MALIGNANT"). These transformations are necessary to input into the pretrained Inception V3 model.

We split the 1317 images into training, validation, and test sets at 70%, 15%, and 15% of the data, respectively.

In building the training model, we loaded the Inception V3 CNN trained on ImageNet, removing the top layer that outputs response data to the previous image categories in order to add our own layers that pool the previous layers, add a dense layer (which makes all the weighted connections), and then creates a softmax that creates a probabilistic response function between our binary responses.

We then froze all other layers, and trained just the new layers to get a baseline of accurate weights. We then unfroze the main top 3 layers of inception modules and trained them on our dataset for 80 epochs, with a patience factor of 30 (meaning if the validation error didn't improve in 30 epochs, the model would stop training).

Through our model iterations we added additional features, including a dropout layer to reduce overtraining, tested different frozen and nonfrozen modules, and augmented our data, in addition to many other features. A large majority and planned features of our model was based on the keras.io repository, and recent research papers³⁻⁹.

For all these models built and trained in the last 48 hours, we compared accuracy on the validation dataset to choose our final prediction model, reaching an 68% diagnostic accuracy rate on the test data.

The final output of the script is the trained inceptionv3_model.h5py deep learning model (around 150 MB, so it could not be uploaded to github, but can be reproduced by running the file)

Convolutional_neural_network_predict.py

The previously mentioned inceptionv3_model.h5py deep learning model file is used by the convolutional_neural_network_predict.py function for predicting all the newly uploaded mammogram images whether the mass lesion is malignant or benign. The prediction function creates a vector consisting of probability scores for the lesion being malignant or benign. If the softmax classifier for being malignant is higher than 0.5 we classify it as cancerous lesion.

The function is also converting the uploaded DICOM image to PNG format so it can be plotted on the UI.

1. <http://marathon.csee.usf.edu/Mammography/Database.html>
2. <https://wiki.cancerimagingarchive.net/display/Public/CBIS-DDSM>
3. Chougrad, et al. Deep Convolutional Neural Networks for Breast Cancer Screening.
<https://www.sciencedirect.com/science/article/pii/S0169260717301451>
4. Qiu Y, Yan S, Tan M, Cheng S, Liu H, Zheng B (2016) Computer-aided classification of mammographic masses using the deep learning technology: a preliminary study:
5. Jiao Z, Gao X, Wang Y, Li J (2016) A deep feature based framework for breast masses classification. Neurocomputing 197:221–231
6. Computational mammography using deep neural networks
7. Probabilistic visual search for masses within mammography images using deep learning
8. https://www.researchgate.net/publication/320321443_Mammogram_Classification_using_Deep_Learning_Features
9. <https://keras.io/applications/>