School of Computer Science Engineering and Information Systems

Fall Semester 2024-25 MTech (Software Engineering)

SWE1011 – Soft Computing – F1+TF1

Project Based Component

REVIEW - 3

THYROID NODULES ANALYSIS THROUGH ULTRASOUND IMAGES USING DEEP LEARNING

Team Members:

MOGITH P – 21MIS0204

MOTHISWAR T B G - 21MIS0258

YUVASHREE D – 21MIS0465

Faculty:

Prof. BHARANI DHARAN N

Introduction:

The thyroid gland, a butterfly-shaped structure near the base of the neck that secretes hormones that are essential for controlling metabolism, growth, and development, can develop abnormal growths or lumps called thyroid nodules. A tiny fraction of thyroid nodules can be cancerous, which poses major health hazards even if the majority are benign. The prevalence of thyroid nodules varies widely across populations, with studies suggesting that they are detectable in up to 68% of individuals through ultrasound imaging, although the majority remain asymptomatic and unnoticed. Since women and the elderly are more likely to be impacted, early detection is especially crucial for these groups.

The possibility of cancer is what gives thyroid nodules their clinical relevance. Differentiating between benign and malignant nodules is essential to ensuring prompt treatment and avoiding needless invasive operations, even though less than 5% of nodules are cancerous. Although fine-needle aspiration biopsy (FNAB) and other conventional techniques for identifying thyroid nodules are successful, they have drawbacks such as invasiveness, patient discomfort, and inconsistent diagnostic precision. Thyroid nodules are frequently evaluated with ultrasound imaging, a non-invasive and extensively accessible technology. It offers crucial structural information and real-time imaging to help assess the risk of cancer. The need for automated and precise diagnostic tools is highlighted by the fact that radiologists' subjective manual interpretation might result in inconsistent diagnoses.

What is ML and DL?

Systems can learn patterns and make predictions or choices based on data without explicit programming thanks to machine learning (ML), a subset of artificial intelligence (AI). Numerous industries, including healthcare, banking, and robotics, heavily rely on machine learning algorithms, including supervised, unsupervised, and reinforcement learning models. A subfield of machine learning called deep learning (DL) uses multi-layered artificial neural networks (ANNs) to process and analyze massive datasets. These networks can learn hierarchical representations of data because they act similarly to the human brain.

By reaching state-of-the-art performance in tasks like image classification, segmentation, and anomaly detection, deep learning models—such Convolutional

Neural Networks (CNNs)—have completely transformed the medical imaging industry. Because CNNs automatically extract elements like edges, forms, and textures—all of which are essential for medical diagnosis—they are very good at interpreting image data. DL models are essential in contemporary healthcare applications such as disease categorization, organ segmentation, and cancer detection because of their capacity to handle complex and high-dimensional data.

There are several advantages to using deep learning for thyroid nodule analysis:

- **Enhanced Accuracy:** DL models offer a greater level of diagnostic dependability and sometimes even surpass skilled radiologists and conventional statistical techniques.

- **Time Efficiency:** Automated technologies minimize diagnostic delays by processing and classifying huge volumes of ultrasound pictures quickly.

- **Non-Invasive Assessment:** DL-based solutions enhance patient comfort and reduce the hazards associated with invasive treatments by decreasing the reliance on biopsies.

- **Standardization:** By eliminating the variability brought about by human interpretation, automated technologies guarantee consistency in diagnosis.

Problem Statement:

Thyroid nodules are a common finding in clinical practice, and accurately distinguishing between benign and malignant nodules is crucial for determining appropriate treatment plans. However, manual interpretation of ultrasound images can be prone to human error, leading to misdiagnosis and potential treatment delays. This highlights the need for an automated, reliable diagnostic solution to assist healthcare professionals. The goal of this project is to develop a deep learning model, specifically a convolutional neural network (CNN), capable of analyzing ultrasound images and classifying thyroid nodules as benign or malignant. By utilizing a large, labeled dataset and employing image preprocessing and augmentation techniques, the model aims to improve classification accuracy. The performance of the model will be evaluated using metrics such as accuracy, sensitivity, and specificity, with the objective of enhancing diagnostic efficiency and ultimately improving patient care.

Objectives:

- **Develop a high-accuracy model for thyroid nodule classification** by leveraging advanced deep learning techniques, such as Convolutional Neural Networks (CNNs) and transfer learning, to effectively analyze ultrasound images and differentiate between benign and malignant nodules.
- **Enhance model reliability and generalizability** through robust preprocessing, integration of clinical data, and comprehensive validation across diverse datasets, ensuring consistent performance in real-world clinical scenarios.

Dataset Description:

The Algerian Ultrasound Images Thyroid Dataset (AUITD) contains thyroid ultrasound images collected from hospitals in city of Setif, Algeria, and was labeled by volunteer doctors. The dataset includes three classes: benign, malignant, and normal. It features 1937 images totally. This dataset is intended for developing machine learning models to classify thyroid nodules and aid in diagnostic tasks. With labeled ultrasound images, the AUITD offers an opportunity to train and validate deep learning models in the medical domain, enhancing the accuracy of thyroid disease diagnosis.

Fundamentals of ML/DL technique used:

Model 1: Ensemble of CNN and LSTM

1. **Convolutional Neural Networks (CNNs):**

   CNNs are highly effective for processing image data. They use convolutional layers to extract spatial features like edges, shapes, and textures from ultrasound images, which are crucial for detecting thyroid nodules. Pooling layers further downsample feature maps, reducing dimensionality while retaining important information.

2. **Long Short-Term Memory (LSTM):**

LSTMs are a type of recurrent neural network (RNN) designed to handle sequential data. In this context, LSTMs process temporal or sequential patterns in the ultrasound imaging data (if captured in sequences) or extracted spatial features to identify contextual dependencies that might improve classification performance.

Model 2: CNN and LSTM with Attention Mechanisms (Spatial and CBAM)

1. **Attention Mechanisms:**

   Attention mechanisms allow the model to focus on the most relevant parts of the input.

   - **Spatial Attention:** Focuses on specific regions in the input image that are more likely to contain nodules.
   - **CBAM (Convolutional Block Attention Module):** Combines channel and spatial attention to refine feature maps, ensuring the model focuses on significant spatial regions and important feature channels.

2. **Integration with CNN and LSTM:**

   Attention mechanisms can be integrated into CNN and LSTM architectures to enhance feature learning by emphasizing relevant areas in the image or sequences.

Model 3: Transfer Learning and Feature-Based Prediction with ML Models

1. **Transfer Learning:**

   Transfer learning involves using pre-trained CNN models (like VGG, ResNet, or Inception) to extract features from images. These models are trained on large datasets (e.g., ImageNet) and can generalize well to similar image classification tasks with minimal fine-tuning.

2. **Feature Extraction:**

   The pre-trained CNN is used as a feature extractor by removing the top classification layers and retaining convolutional layers to output feature vectors.

3. **Machine Learning Models:**

   Traditional ML algorithms like KNN, Random Forest, and SVM are employed to classify nodules based on the extracted features.

   - **K-Nearest Neighbors (KNN):** A simple, non-parametric model that predicts based on feature proximity in the feature space.
   - **Random Forest:** An ensemble learning method that uses multiple decision trees for robust classification.
   - **Support Vector Machine (SVM):** A robust classifier that finds the hyperplane maximizing class separation in the feature space.

Methodology:

1. Data Collection and Labelling:

- Gather a dataset of Algerian thyroid ultrasound images.
- **Details:**
  - Acquire thyroid ultrasound images from hospitals, medical repositories, or publicly available datasets.
  - Manually label the images with appropriate disease classifications (e.g., healthy, benign, malignant).

2. Data Preprocessing:

- Prepare the collected data for model training.
- **Steps:**
  - Normalize image intensities to ensure consistency.
  - Resize images to a uniform shape compatible with the model.
  - Remove noise and artifacts using image denoising techniques.

    - Convert images to grayscale or RGB format as required.

3. Apply Data Augmentation:

- Enhance dataset variability to prevent overfitting.
- **Techniques:**
    - Rotate images at various angles.
    - Flip, zoom, or crop regions of interest.
    - Apply transformations like brightness adjustment or adding Gaussian noise.

4. Split Train-Validation-Test

- Divide the dataset for model training, validation, and testing.
- **Ratios:**
    - Typically, split as 70% training, 15% validation, and 15% testing.
- **Purpose:**
    - Train the model on the training set.
    - Validate hyperparameters and detect overfitting on the validation set.
    - Evaluate final model performance on the test set.

5. Model Selection

- Choose the appropriate models for experimentation.
- **Models:**
    - **RNN (Recurrent Neural Network):** If time-sequential features (e.g., ultrasound image sequences) are important.

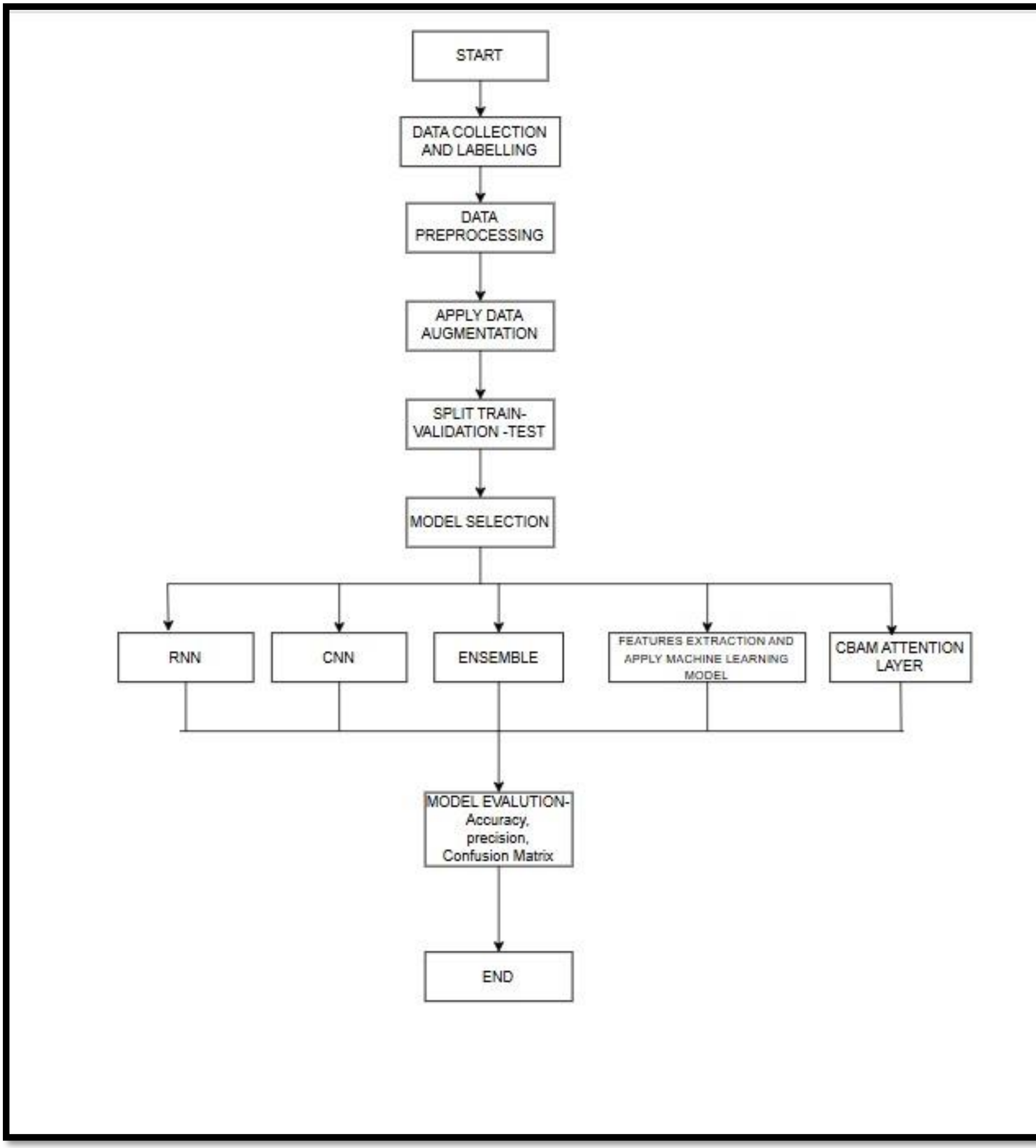
- **CNN (Convolutional Neural Network):** For spatial feature extraction in thyroid ultrasound images.
- **Ensemble Models:** Combine predictions of multiple models (e.g., RNN + CNN) to improve accuracy.
- **Feature Extraction and Apply Machine Learning:** Use pre-trained feature extractors (e.g., VGG, ResNet) and train machine learning models like SVM or Random Forest on these extracted features.
- **CBAM (Convolutional Block Attention Module) Attention Layer:** Enhance CNN by focusing on critical regions of the ultrasound images to improve disease localization.

6. Model Evaluation:

- **Objective:** Measure model performance using metrics.
- **Metrics:**
    - **Accuracy:** Percentage of correct predictions.
    - **Precision:** Ability of the model to correctly identify positive cases.
    - **Confusion Matrix:** Evaluate true positives, false positives, true negatives, and false negatives for deeper insights into model performance.

7. End

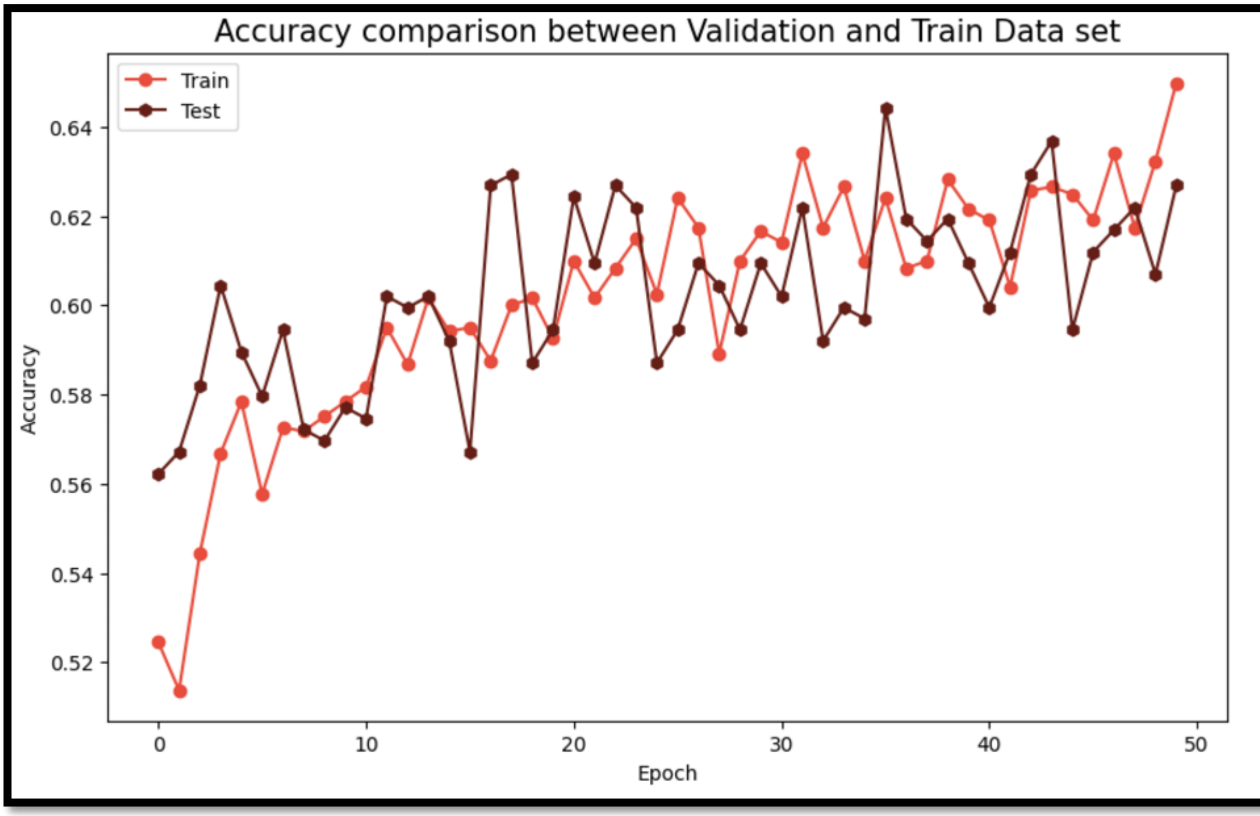
- **Next Steps:**
    - Deploy the best-performing model for practical use (e.g., in a diagnostic tool for Algerian thyroid diseases).
    - Optimize further based on real-world feedback.


START
DATA COLLECTION AND LABELLING
DATA PREPROCESSING
APPLY DATA AUGMENTATION
SPLIT TRAIN-VALIDATION -TEST
MODEL SELECTION
RNN
CNN
ENSEMBLE
FEATURES EXTRACTION AND APPLY MACHINE LEARNING MODEL
CBAM ATTENTION LAYER
MODEL EVALUTION-Accuracy, precision, Confusion Matrix
END

Results:

MODEL – 1:

CNN:

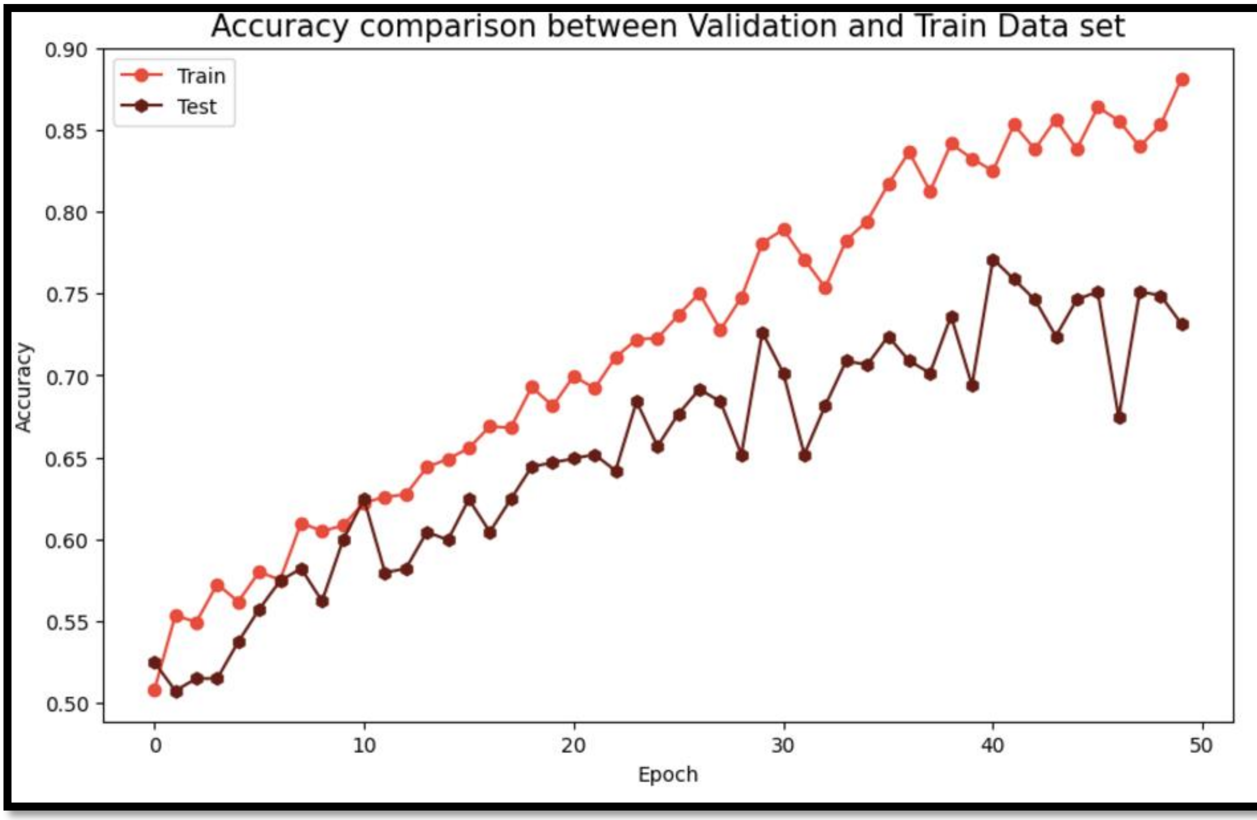



```
13/13 ━━━━━━━━━━━━━━━━━━━━ 1s 97ms/step - accuracy: 0.6147 - loss: 0.7448
Testing Loss: 0.7482492923736572
Testing Accuracy: 0.6268656849861145
13/13 ━━━━━━━━━━━━━━━━━━━━ 2s 103ms/step
                precision    recall  f1-score   support

        Benign       0.55      0.78      0.64       152
     Malignant       0.75      0.66      0.70       208
normal thyroid       0.40      0.05      0.09        42

      accuracy                           0.64       402
     macro avg       0.57      0.49      0.48       402
  weighted avg       0.64      0.64      0.62       402
```

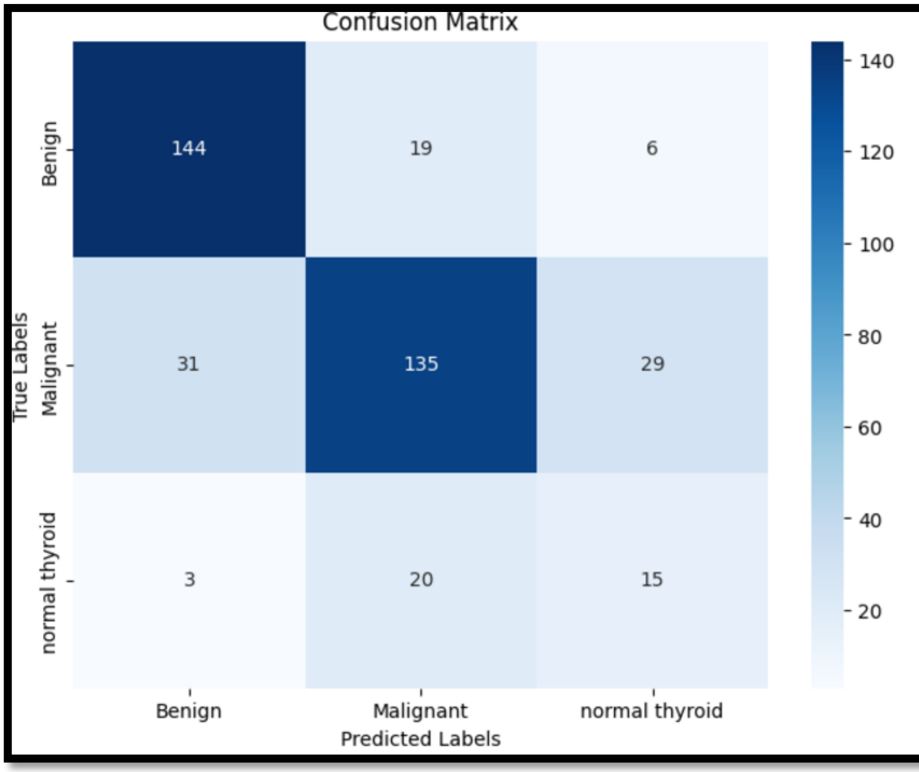
LSTM:




```
13/13 ━━━━━━━━━━━━━━━━━━━━ 0s 3ms/step - accuracy: 0.7237 - loss: 0.7977
Testing Loss: 0.7773849964141846
Testing Accuracy: 0.7313432693481445
13/13 ━━━━━━━━━━━━━━━━━━━━ 0s 22ms/step
              precision    recall  f1-score   support

           0       0.81      0.85      0.83       169
           1       0.78      0.69      0.73       195
           2       0.30      0.39      0.34        38

    accuracy                           0.73       402
   macro avg       0.63      0.65      0.63       402
weighted avg       0.74      0.73      0.74       402
```

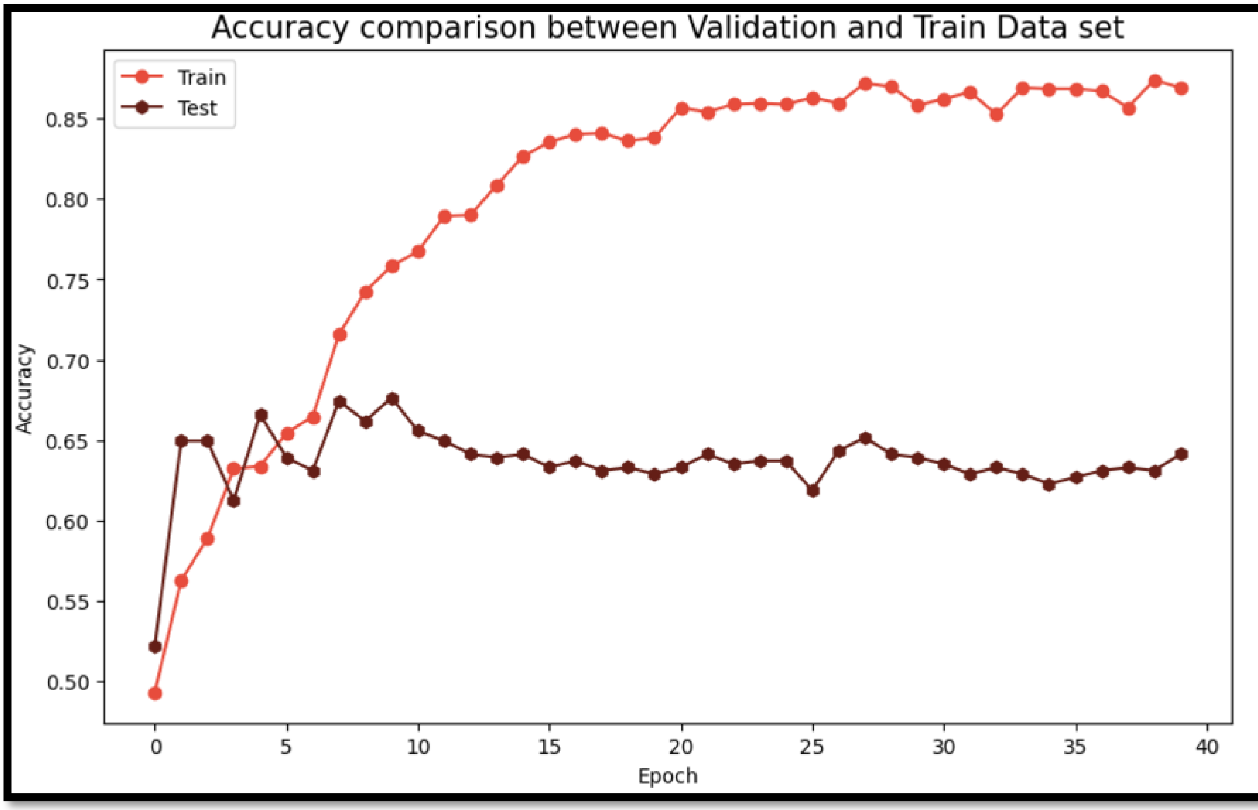


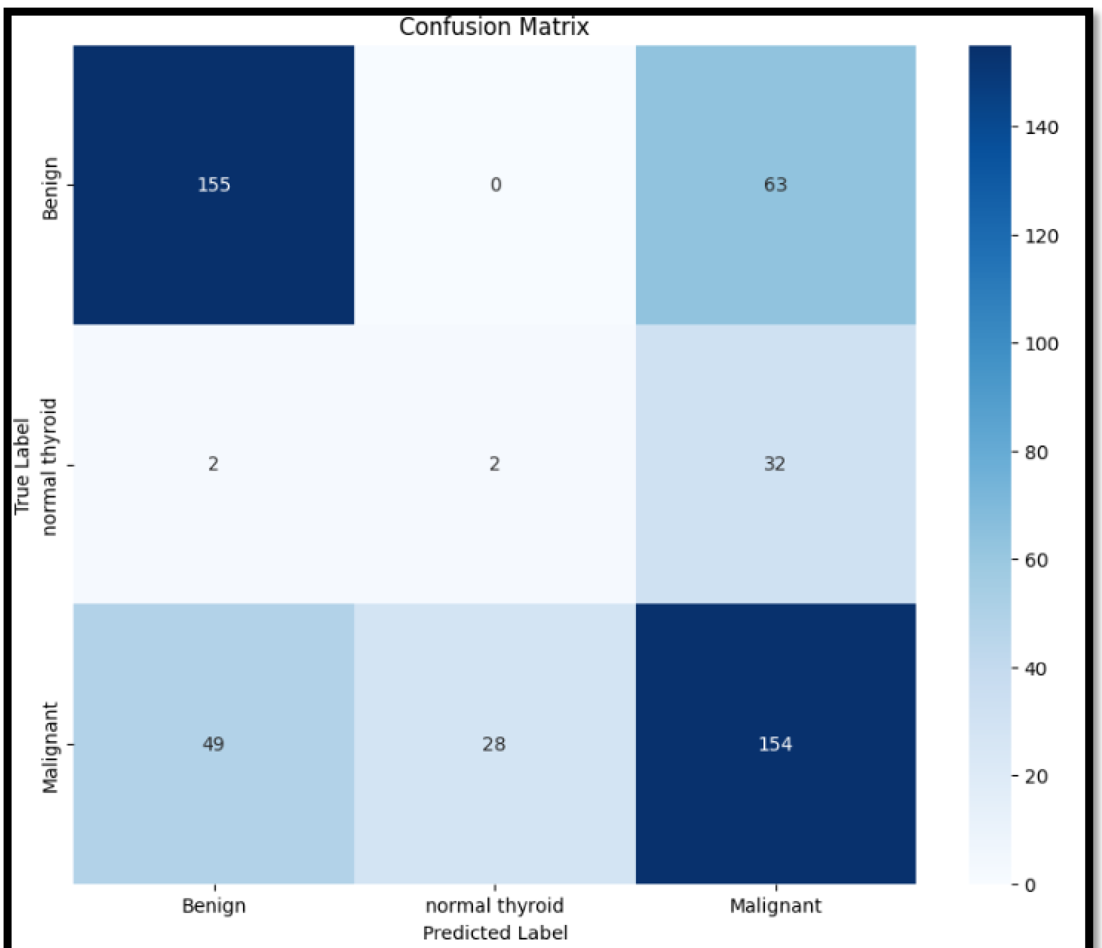
ENSEMBLE:

```
Found 1937 files belonging to 3 classes.
Prepared CNN and LSTM data.
```

Model: "functional_13"

Layer (type)	Output Shape	Param #	Connected to
input_layer_2 (InputLayer)	(None, 50, 50, 3)	0	-
conv2d_2 (Conv2D)	(None, 48, 48, 32)	896	input_layer_2[0]…
max_pooling2d_2 (MaxPooling2D)	(None, 24, 24, 32)	0	conv2d_2[0][0]
conv2d_3 (Conv2D)	(None, 22, 22, 64)	18,496	max_pooling2d_2[…
input_layer_3 (InputLayer)	(None, 28, 28, 3)	0	-
max_pooling2d_3 (MaxPooling2D)	(None, 11, 11, 64)	0	conv2d_3[0][0]
time_distributed_1 (TimeDistributed)	(None, 28, 84)	0	input_layer_3[0]…
flatten_2 (Flatten)	(None, 7744)	0	max_pooling2d_3[…
lstm_1 (LSTM)	(None, 128)	109,056	time_distributed…
concatenate (Concatenate)	(None, 7872)	0	flatten_2[0][0], lstm_1[0][0]
dense_4 (Dense)	(None, 128)	1,007,744	concatenate[0][0]
dropout_1 (Dropout)	(None, 128)	0	dense_4[0][0]
dense_5 (Dense)	(None, 3)	387	dropout_1[0][0]

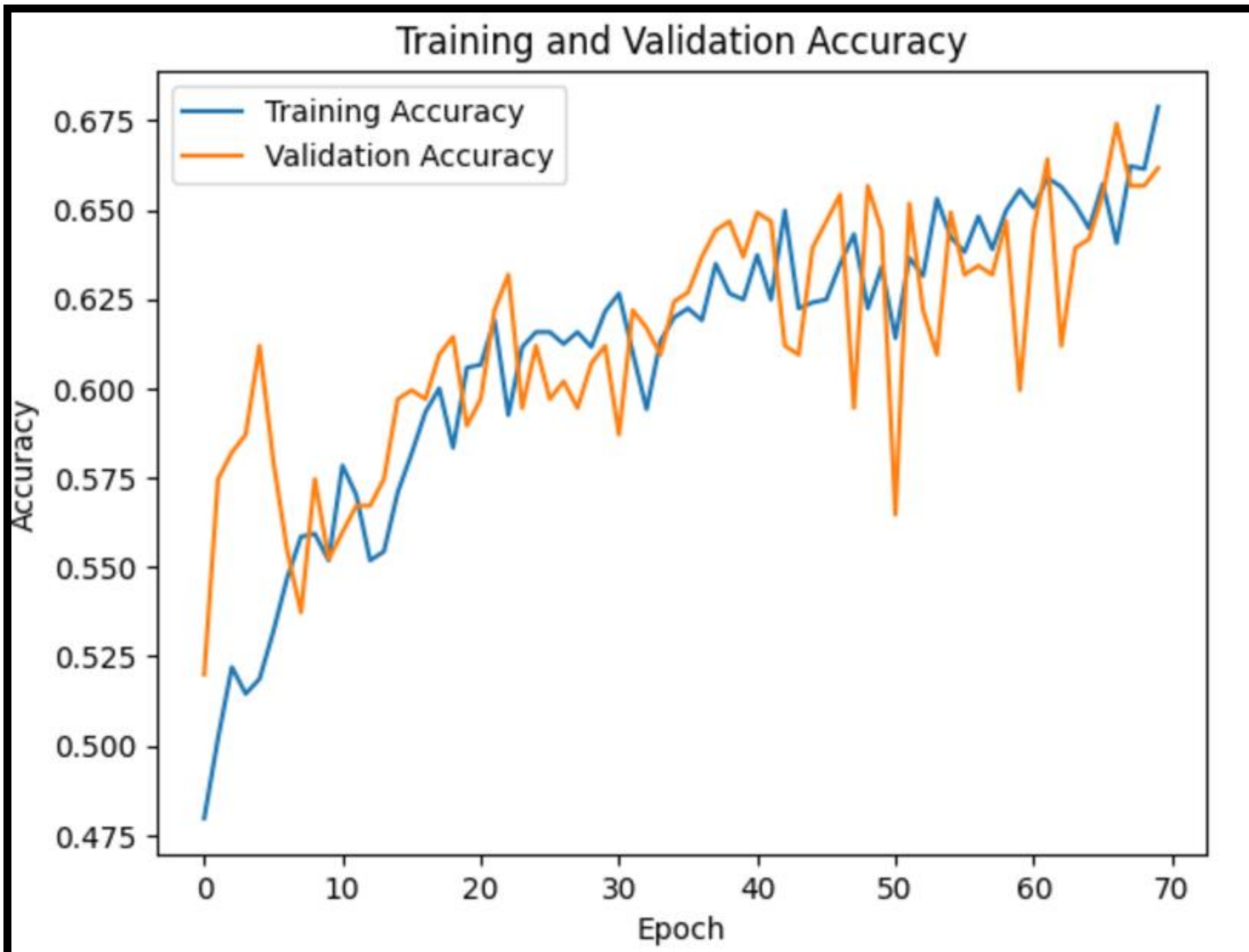

Accuracy comparison between Validation and Train Data set
Train
Test
Accuracy
0.85
0.80
0.75
0.70
0.65
0.60
0.55
0.50
0
5
10
15
20
25
30
35
40
Epoch


Confusion Matrix
True Label
Benign
normal thyroid
Malignant
155
0
63
2
2
32
49
28
154
140
120
100
80
60
40
20
0
Benign
normal thyroid
Malignant
Predicted Label

MODEL – 2:

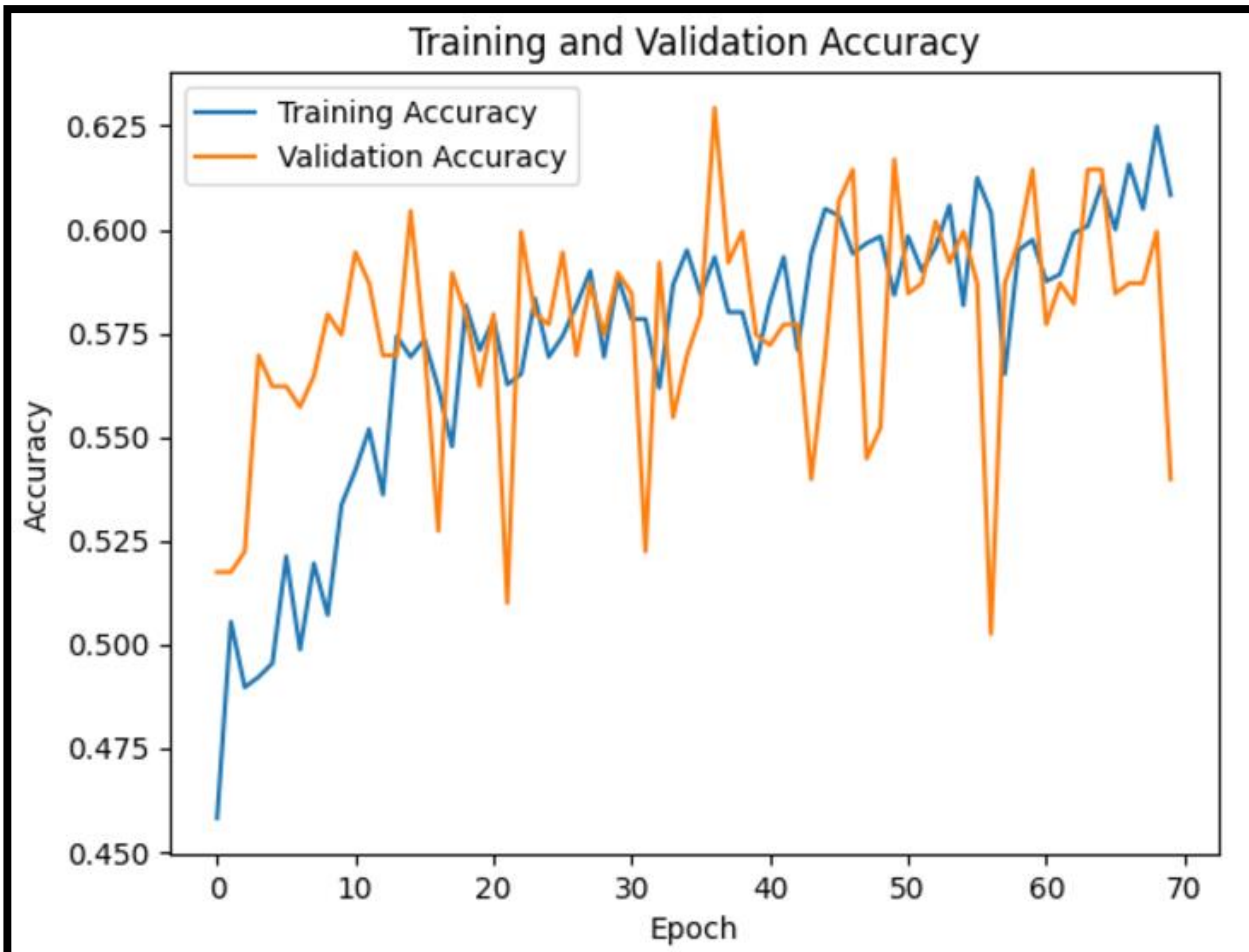
CBAM (Spatial, Channel) **with CNN:**

```
38/38 ━━━━━━━━━━━━━━━━━━━━ 5s 119ms/step - accuracy: 0.6517 - loss: 0.7198 - val_accuracy: 0.6542 - val
_loss: 0.7457
Epoch 67/70
38/38 ━━━━━━━━━━━━━━━━━━━━ 5s 114ms/step - accuracy: 0.6471 - loss: 0.7295 - val_accuracy: 0.6741 - val
_loss: 0.7101
Epoch 68/70
38/38 ━━━━━━━━━━━━━━━━━━━━ 5s 115ms/step - accuracy: 0.6891 - loss: 0.6907 - val_accuracy: 0.6567 - val
_loss: 0.7183
Epoch 69/70
38/38 ━━━━━━━━━━━━━━━━━━━━ 5s 115ms/step - accuracy: 0.6435 - loss: 0.7343 - val_accuracy: 0.6567 - val
_loss: 0.7449
Epoch 70/70
38/38 ━━━━━━━━━━━━━━━━━━━━ 5s 116ms/step - accuracy: 0.6866 - loss: 0.7000 - val_accuracy: 0.6617 - val
_loss: 0.7014
13/13 ━━━━━━━━━━━━━━━━━━━━ 1s 99ms/step - accuracy: 0.6751 - loss: 0.7279
Training Accuracy: 67.88%
Validation Accuracy: 66.17%
Test Accuracy: 64.43%
```



CBAM with LSTM:

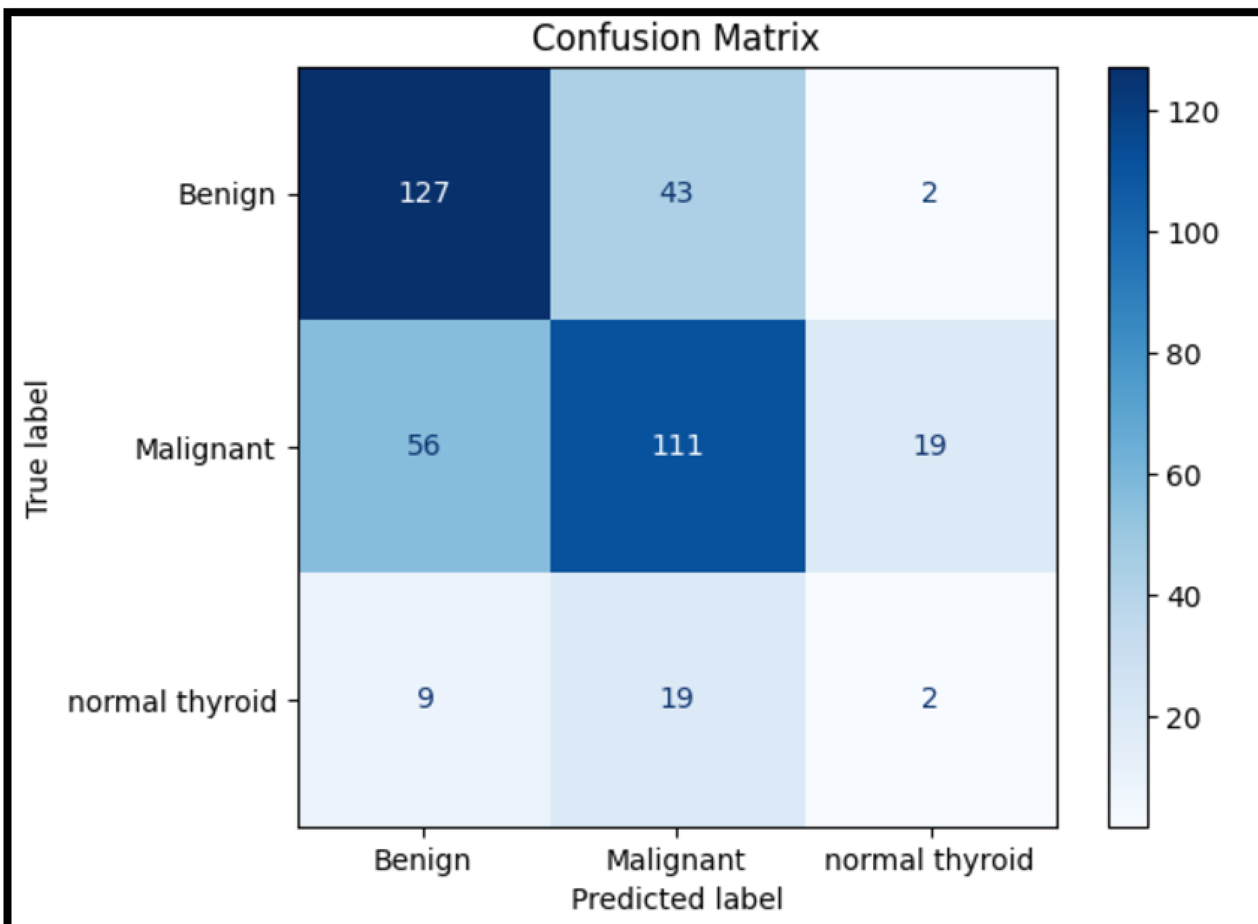
```
38/38 ━━━━━━━━━━━━━━━━━━━━ 5s 117ms/step - accuracy: 0.6248 - loss: 0.7810 - val_accuracy: 0.6144 - val_loss: 0.7933
Epoch 65/70
38/38 ━━━━━━━━━━━━━━━━━━━━ 5s 115ms/step - accuracy: 0.6072 - loss: 0.7647 - val_accuracy: 0.6144 - val_loss: 0.7988
Epoch 66/70
38/38 ━━━━━━━━━━━━━━━━━━━━ 5s 117ms/step - accuracy: 0.6014 - loss: 0.7782 - val_accuracy: 0.5846 - val_loss: 0.7863
Epoch 67/70
38/38 ━━━━━━━━━━━━━━━━━━━━ 5s 117ms/step - accuracy: 0.6190 - loss: 0.7780 - val_accuracy: 0.5871 - val_loss: 0.8152
Epoch 68/70
38/38 ━━━━━━━━━━━━━━━━━━━━ 5s 121ms/step - accuracy: 0.6206 - loss: 0.7558 - val_accuracy: 0.5871 - val_loss: 0.7905
Epoch 69/70
38/38 ━━━━━━━━━━━━━━━━━━━━ 5s 120ms/step - accuracy: 0.6114 - loss: 0.7873 - val_accuracy: 0.5995 - val_loss: 0.7800
Epoch 70/70
38/38 ━━━━━━━━━━━━━━━━━━━━ 5s 119ms/step - accuracy: 0.6111 - loss: 0.7639 - val_accuracy: 0.5398 - val_loss: 0.8209
13/13 ━━━━━━━━━━━━━━━━━━━━ 1s 97ms/step - accuracy: 0.5134 - loss: 0.8396
Training Accuracy: 60.83%
Validation Accuracy: 53.98%
Test Accuracy: 54.48%
```



MODEL – 3:

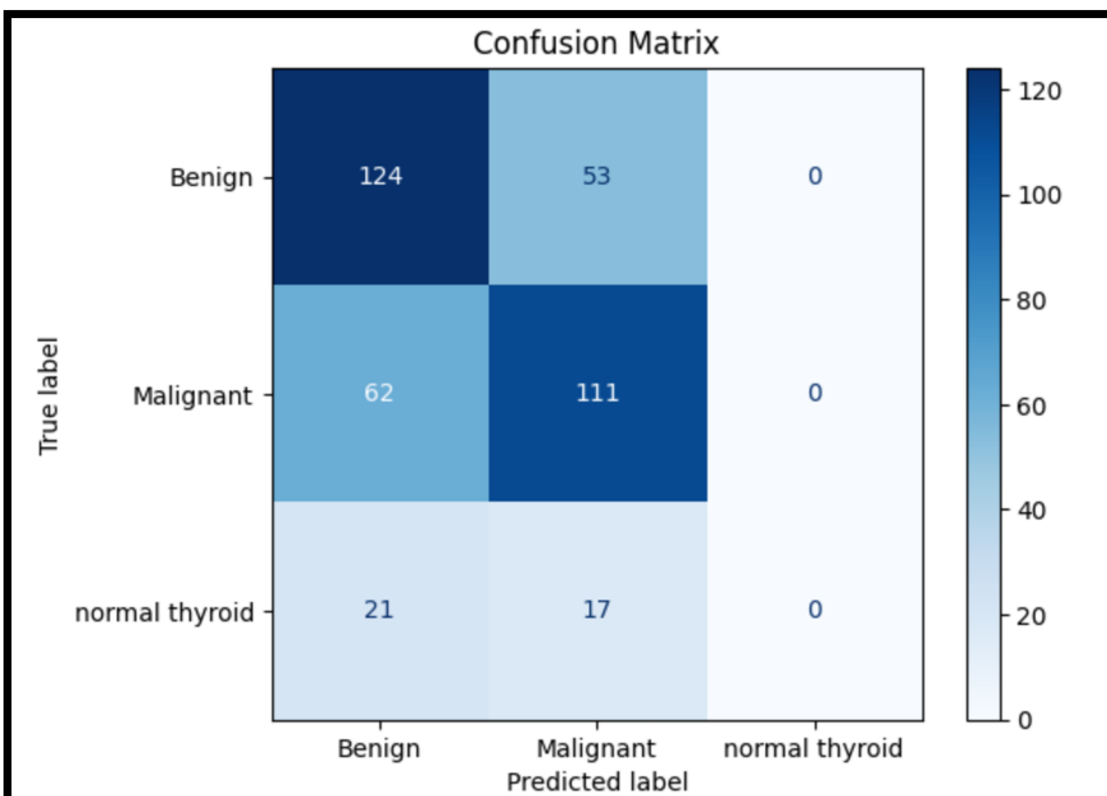
VGG with KNN:

```
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 85ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 81ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 99ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 84ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 87ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 80ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 85ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 96ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 78ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 89ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 85ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 108ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 461ms/step
Train Accuracy: 0.7779
Test Accuracy: 0.6186
Validation Accuracy: 0.6186
```



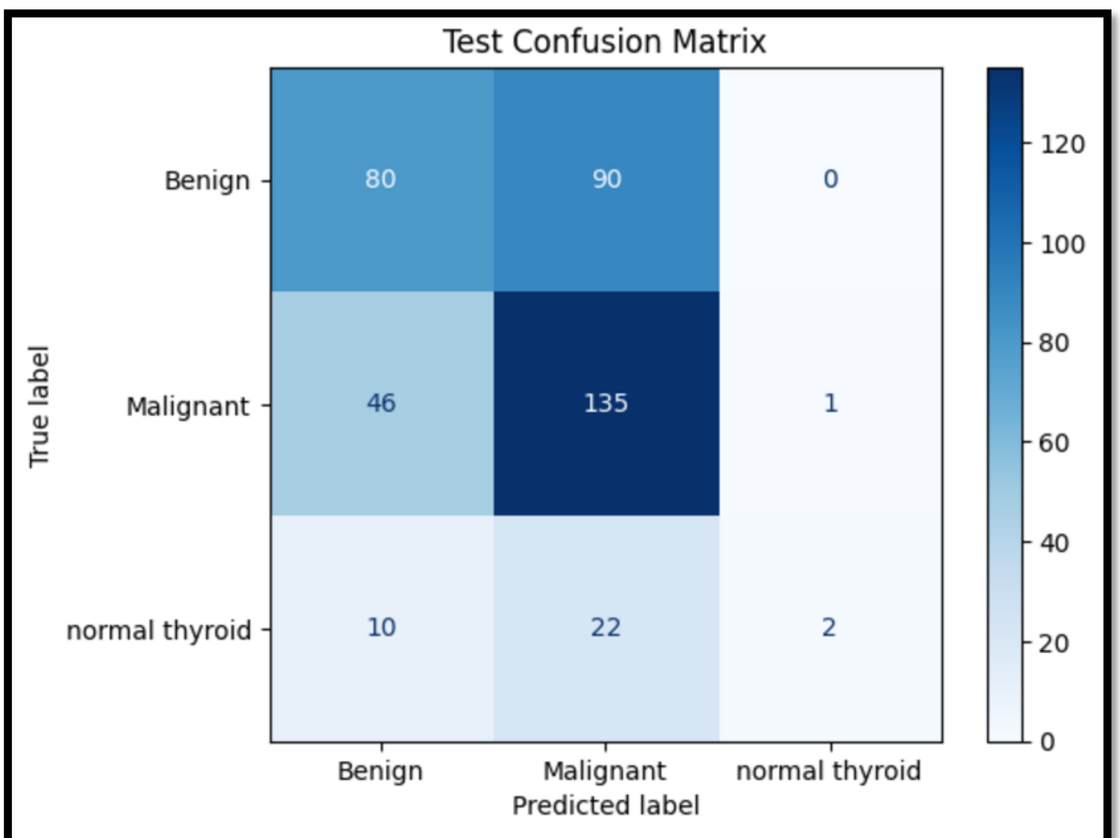
VGG with SVM:

```
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 26ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 26ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 37ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 26ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 26ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 26ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 26ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 26ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 26ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 27ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 26ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 26ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 26ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 35ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 26ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 26ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 26ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 26ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 26ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 27ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 26ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 4s 4s/step
Train Accuracy: 0.6533
Test Accuracy: 0.6057
Validation Accuracy: 0.6057
```



ResNet50 with KNN:

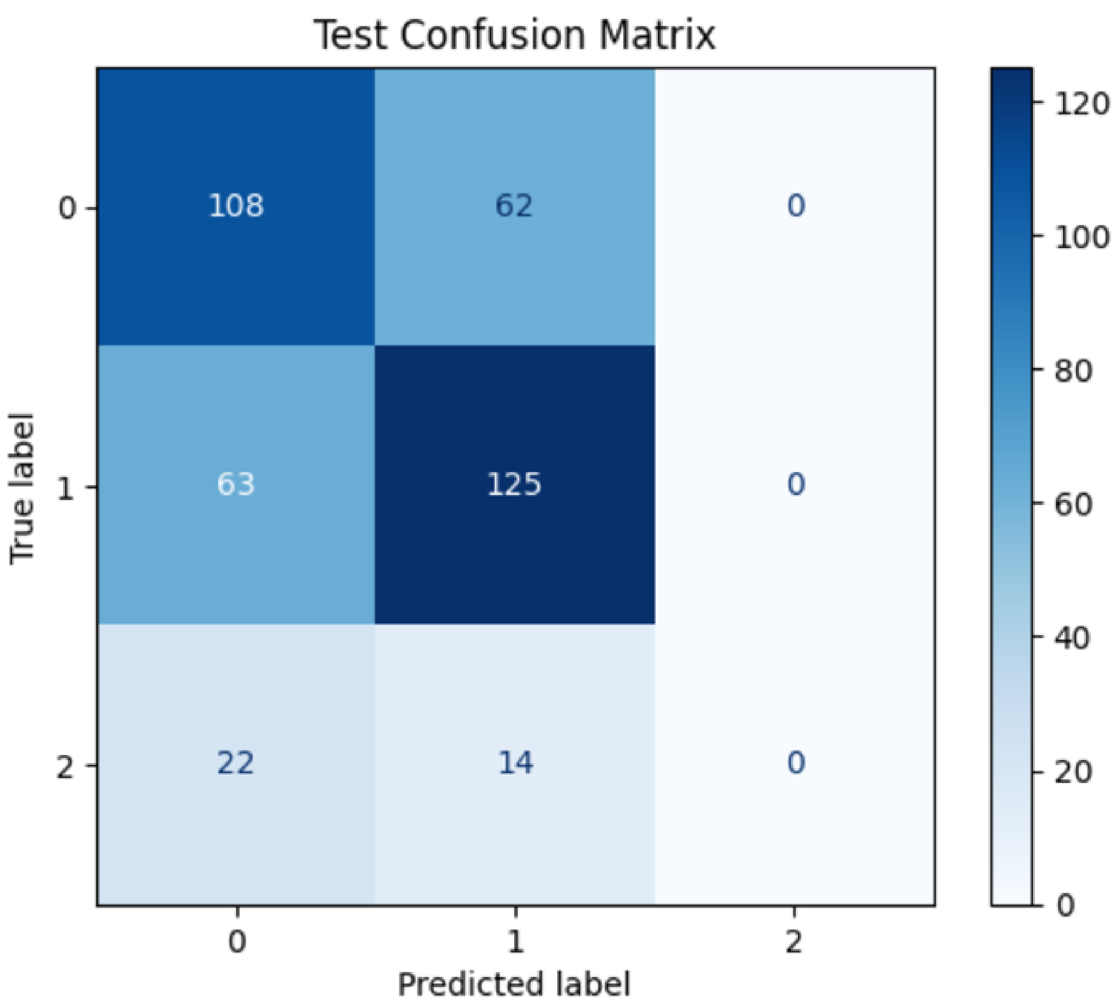
```
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 29ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 1s 1s/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 38ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 29ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 31ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 29ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 30ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 30ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 31ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 29ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 29ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 30ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 30ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 30ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 22ms/step
Train Accuracy: 0.7395
Validation Accuracy: 0.5622
Test Accuracy: 0.5622
```



ResNet50 with SVM:

```
Found 1937 images belonging to 3 classes.
40/40 ━━━━━━━━━━━━━━━━━━━━ 8s 110ms/step
10/10 ━━━━━━━━━━━━━━━━━━━━ 1s 153ms/step
13/13 ━━━━━━━━━━━━━━━━━━━━ 2s 125ms/step
Train Accuracy: 0.6365
Validation Accuracy: 0.5841
Test Accuracy: 0.5914
```

Test Confusion Matrix:



	Training Accuracy	Testing Accuracy	Precision	Recall	F1-score
CNN	66.44	62.68	0.64	0.65	0.61
LSTM	89.05	73.13	0.75	0.75	0.75
Ensemble	87.87	64.12	-	-	-
CBAM with CNN	67.88	64.43	-	-	-
CBAM with LSTM	60.83	53.98	-	-	-
VGG with KNN	77.79	61.86	-	-	-
VGG with SVM	65.33	60.57	-	-	-
ResNet with KNN	73.95	56.22	-	-	-
ResNet with SVM	63.65	59.14	-	-	-

Conclusion:

In the study of Algerian ultrasound thyroid disease prediction, various models were trained and evaluated. Among these, the LSTM model achieved a test accuracy of 73.13%, while the ensemble method followed with an accuracy of 64.12%, outperforming other approaches. However, it was observed that all models suffered from overfitting, primarily due to the significant class imbalance in the dataset.

To address this limitation, future work should focus on balancing the dataset to enhance model generalization. Techniques such as Generative Adversarial Networks (GANs) and Synthetic Minority Oversampling Technique (SMOTE) can be employed to generate synthetic data and mitigate class imbalance. Once the dataset is balanced, the current models can be retrained to potentially achieve better performance and robustness in predicting thyroid disease from ultrasound images.

References:

1. https://www.geeksforgeeks.org/convolutional-neural-network-cnn-in-machine-learning/

2. https://www.geeksforgeeks.org/deep-learning-introduction-to-long-short-term-memory/

3. https://www.ai-contentlab.com/2022/12/introduction-to-cnns-with-attention.html

4. https://www.digitalocean.com/community/tutorials/attention-mechanisms-in-computer-vision-cbam