Proposed approach of the study:
Classification approach was best suited for predicting poverty using satellite imagery.
Classifier: Convolutional Neural Network is applied to extract features such as edges, shapes, corners, and pixel intensities. The output of the trained CNN model will be one of the four classes – agriculture substances, building roof, road, water.

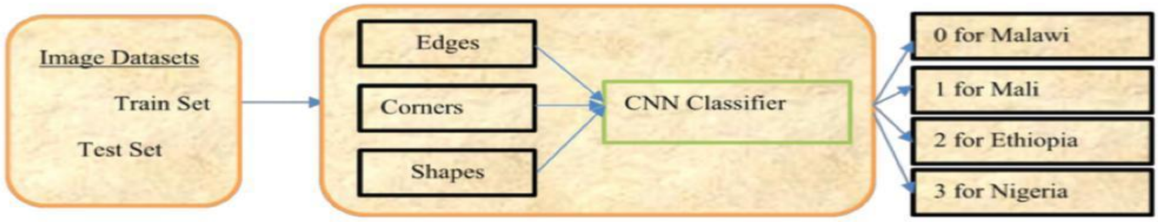

Fig. 1 Showing CNN Classifier [8]

The study model is implemented in two pre-trained models (ResNet50, VGG16) and a developed CNN model from the beginning to classify/segment the high-resolution satellite pictures into four groups of classes as four ranks of poverty: very high, high, medium, and low poverty.
By feeding each model with the augmented images, then the CNNs algorithm will automatically understand to extract many features that recognize some primary factors. Such as roads, buildings, vegetation and & farmlands, water resources, and building roof, discover relative image patterns and make predictions of the livelihood well-being or poverty level.
The data preprocessing techniques was with the help of the Keras library frame on top of TensorFlow using a python programming language. This technique allows for the easy execution of the architectural neural network, i.e., convolutional neural network (CNN) for image segmentation. The input layers accept in standard image size as already stated there rescaling regular of occurrence are divided by 255 across the dataset to make easy accessibility of the neural network. It has an import layer acceptable for RGB picture height, width, and depth. Eight (8) batch sizes were initialized as datasets generators' sizes. There was an adjustment of the size of the training set in adjusting the model overfitting and enhancing the generalization.

1. **CNN-**
   The study describes a neural network architecture with 8 convolutional layers, 2 connected layers, and a SoftMax output layer. Convolutional layers use ReLU activation, and batch normalization. Max pooling and drop-out are applied for regularization. The final output layer categorizes data into four classes using SoftMax for multi-class segmentation problems.
   Experimentation- Adam optimization ie-4 (0.0001) was also applied as the learning rate with the decay of ie-4. They are all optimized by the learning rate by the number of epochs defined. As the class was multiple groups, a categorical cross-entropy was applied since it is not in binary mode. Early stopping was initialized into the model to monitor overfitting, and image batch sizes of eight (8) were set as model generators. Both the train and validation images were passed to the network as generators.
2. **VGG16-**
   VGG16 is a 16-layer pre-trained CNN architecture designed for image classification tasks. It takes 256x256x3 images as input with weights from ImageNet. It uses pairs of convolution layers with 4x4 average pooling, followed by flattening and two

4096-node dense layers with ReLU activation and 0.5 dropout. The output is four classes, handled by SoftMax activation. Modifications include adjusting frozen layers and applying global average pooling to reduce overfitting.
<u>Experimentation-</u> VGG16 is a pre-train CNN architecture that consists of 16 layers, and the total Epochs of 200 were applied with a learning rate set as 1e-4(0.001). Adam optimization 1e-4 was used as the learning rate with the decay of 1e-4. All the optimization by the learning rate is divided by the total number of epochs defined. The epoch was reduced to 100 during the experiment but was stopped at 16 epochs. The early stopping technique was initialized into the model to monitor overfitting, and image batch sizes of eight (8) were set as model generators. The VGG network training continues from 96 epochs to 100. As the class was multiple groups, a categorical cross-entropy was applied since it is not in binary mode. Hyperparameters rates for every network model were regarded as the same technique. Both the train and validation images were passed to the network as generators.

3. RESNET50-
   The ResNet50 architecture is a deep neural network designed to address the vanishing gradient problem. It starts with a 256x256 input, uses 3x3 convolutional layers, 7x7 average pooling, flattening, and then three dense layers with ReLU activation (4096, 2048, 1024 nodes) and 0.5 dropout. The output is handled by a SoftMax activation.
   <u>Experimentation-</u> Epochs of 200 were initialized with the learning rate of 1e-4 (0.0001), Adam optimization with learning and decay of 1e-4 are divided by the epochs defined. The callback monitors were used then adjusted the learning rate by remaking it 1e-5 (0.00001). Also, the same with the learning rate and the decay rate is divided by the epochs defined. Categorical crossentropy was applied to maintain the multiple class. The model was trained using 200 epochs, and the early stopping method was used; as a result, the Resnet50 model stopped after executing 59 epochs.

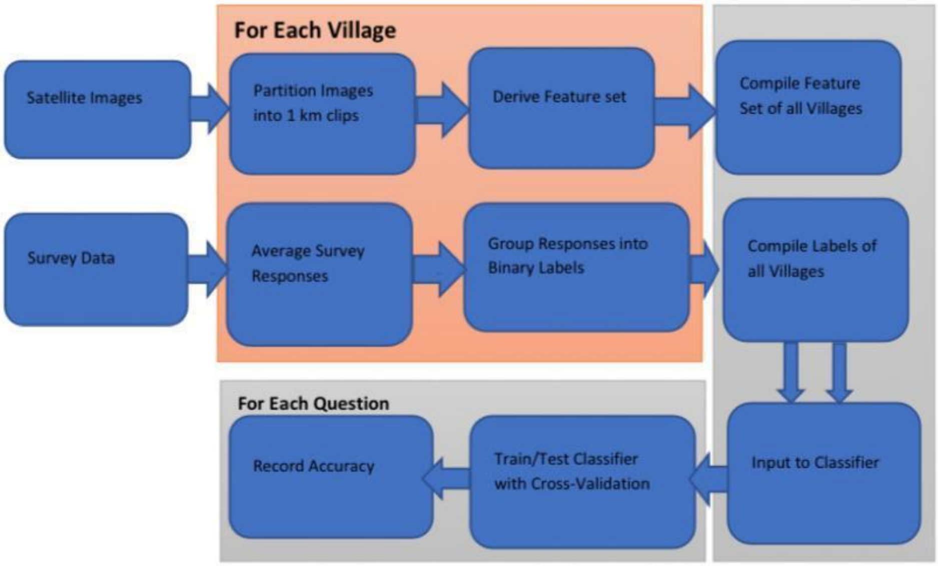

Fig. 2 Showing CNN satellite image and survey Procedure

TABLE II. SHOWING THE MODELS PARAMETERS

Hyper-Parameters	Models		
	CNN	VGG16	ResNet50
Learning-Rate	1e-4	1e-4	1e-4
Drop-out	0.5	0.5	0.5
Hidden-layer activation function	RELU	RELU	RELU
Output-layer activation layer	SoftMax	SoftMax	SoftMax
Epochs	200	200	200
Size of batch in training	8	8	8
Size of batch for validation	8	8	8

Table II. above shows the applied parameters while running the algorithm or model from the learning rate, drop-out, hidden layer, output layer, activation function, etc.

FINDINGS-

Figure3 shows the results that CNN linear model does okay on the training data with 91% accuracy as the accuracy from validation is 71%. The training and validation accuracies are shown in the Figures below.

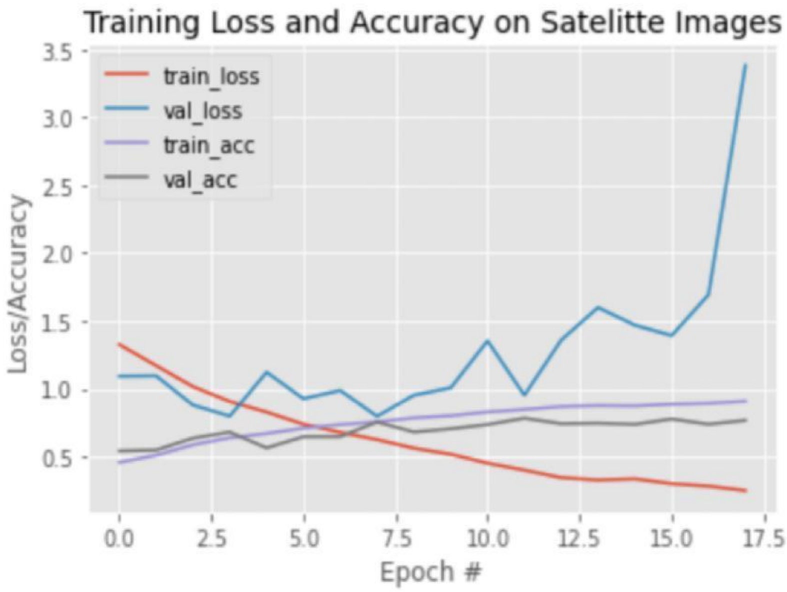

Fig. 3 Showing the performance of CNN model loss and accuracy.

The training loss curves are down while the validation loss is inverse to the training loss in the CNN model, as shown in Figure 3. In terms of VGG16 shown in figure 4, the training accuracy of 94%, while its validation accuracy is 86%. The training and validation accuracies are shown in the Figures below.

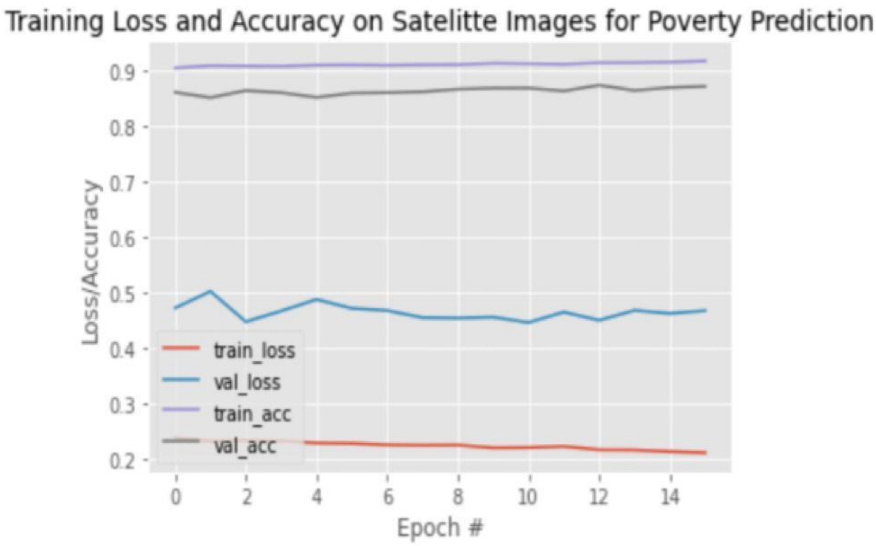

Fig. 4 Showing the performance of VGG16 model loss and accuracy.

The VGG16 also has variations between the training and validation losses, as shown in Figure 4 above. The ResNet50 model in figure 5 shows that the training accuracy is 62% while its validation accuracy of 58%. The training and validation accuracies are shown in the Figures below.

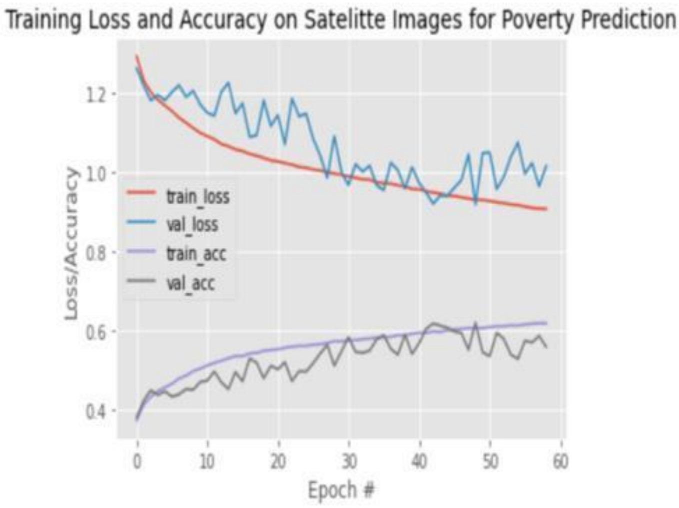

Fig. 5 Showing the performance of ResNet model loss and accuracy.

The ResNet50 model shows that the training and validation losses are on a closed curve, as shown in Figure 5.

TABLE III. SHOWING THE CNN MODEL RESULTS

Countries	Precision	Recall	F1-Score	Support
Nigeria	0.70	0.72	0.71	1154
Mali	0.83	0.64	0.72	1475
Malawi	0.74	0.82	0.78	1270
Ethiopia	0.75	0.83	0.79	1717
Accuracy			0.75	5616
Macro Average	0.76	0.75	0.75	5616
Weighted Average	0.76	0.75	0.75	5616

Table III. represent that the CNN model performs averagely well with an accuracy of 0.75. It also shows that Ethiopia has the highest F1 score while Nigeria has the lowest, with 0.71. CNN performance is better than the ResNet model.

TABLE IV. SHOWING THE VGG16 MODEL RESULTS

Countries	Precision	Recall	F1-Score	Support
Nigeria	0.87	0.77	0.82	1154
Mali	0.81	0.91	0.86	1475
Malawi	0.88	0.88	0.88	1270
Ethiopia	0.93	0.92	0.92	1717
Accuracy			0.87	5616
Macro Average	0.87	0.87	0.87	5616
Weighted Average	0.88	0.87	0.87	5616

The above Table IV. shows that the VGG16 has a performance accuracy of 0.87 which means it performs better than the two models. The country with the lowest F1-score of 0.82 while the country with the highest F1-score of 0.92. The table shows that the VGG16 model performs better than all other models.

TABLE V. SHOWING THE RESNET MODEL RESULTS

Countries	Precision	Recall	F1-Score	Support
Nigeria	0.53	0.59	0.56	1154
Mali	0.64	0.65	0.64	1475
Malawi	0.56	0.59	0.57	1270
Ethiopia	0.74	0.66	0.70	1717
Accuracy			0.62	5616
Macro Average	0.62	0.62	0.62	5616
Weighted Average	0.63	0.62	0.63	5616

The above Table V shows that the ResNet performed weak with an accuracy of 0.62. ResNet from the previous work of other authors shows that the model is not suitable for such kind of analysis.

These weighted average were determined by multiplying every patterns prediction by the algorithm’s model weight to produce the weighted total, then by dividing the value by the sum of the weights. The macro average was calculated straightforward using the normal averaging methods. The macro average F1 score is determined using the arithmetic mean (sometimes known as unweighted mean) of all the per class F1 scores. The countries were classified where zero was assigned to Malawi as a country with the lowest economic status, one was assigned to Mali as a country with a medium economic status, two was assigned to Ethiopia as the country with the high economic status, while three was assigned to Nigeria as the country with the very high economic status. The classification was applied using the convolutional neural network, one of the vital models of deep learning techniques.