Deep Learning-Based Identification of Dermatological Issues in Facial Skin

Alisultan Amanqos*, BSCS Candidate; **Batyr Kenzheakhmetov***, BSCS Candidate; **CJ Chung**, PhD, Professor

College of Arts & Science, Lawrence Technological University
(*)College of Informational Technologies, Astana IT University

INTRODUCTION and ABSTRACT

The project revolves around employing deep learning to automatically identify and classify different types of acne lesions on people's faces. Acne, a common skin condition, presents various lesions, including comedones, nodules, papules, and pustules. A comedone is a type of acne lesion caused by a clogged hair follicle in the skin, while nodules are larger, solid, dome-shaped, or irregularly shaped lesions that develop deeper within the skin. Nodular acne is considered severe and can lead to scarring if not treated properly. Papules are small, raised bumps that typically appear on the skin's surface. They are usually red and inflamed and do not contain pus. Papules are a common form of inflammatory acne. Finally, pustules are small, inflamed lesions filled with pus. Pustules are often tender to the touch and can be a symptom of moderate to severe acne.

By leveraging deep learning algorithms, we aim to develop a smart system capable of accurately recognizing these types of acne lesions. It is important to note that this task is a multi-label classification problem since several types of acne can be found on a person's face, therefore a special model is required for such tasks. The main CNN model achieved 82% accuracy on the test using BatchNormalization, Dropout technologies, and a 5x5 kernel size[6]. Other models using the pre-trained technique showed poor loss results, reaching values greater than one (This includes pre-trained models such as EfficientNetB3, ResNet50, and VGG19).

The prospects for this technology are incredibly promising, with the potential to revolutionize how acne is diagnosed and treated. As research progresses and datasets expand, the accuracy of these systems is likely to improve even further. This means more precise and reliable diagnoses, leading to better treatment outcomes for individuals suffering from acne. Also, It may become possible to remotely monitor the progression of acne lesions through photographs or live video feeds. This could facilitate telemedicine consultations, allowing dermatologists to assess patients' skin conditions without requiring in-person visits.

DATA COLLECTION

Data collection method:
To build our model, we used a free dataset that we could find on the Roboflow website[8]. All pimples in the images have already been classified. To do this, in addition to the images, we were provided with a text document that contained information about each image:

levle1_596_png_jpg.rf.008eb6e34ffe44abfb8d5e1ef588ef52.jpg 127,68,143,84,0 117,139,134,154,2

The Name of the picture - levle1_596_png_jpg.rf.008eb6e34ffe44abfb8d5e1ef588ef52.jpg
Location of a pimple by pixels - *127,68,143,84*
Classification of pimple - *0*

A similar text file was contained in each of the three directories (train, validation, and test). The number of images in the folders is as follows: In the train directory - 3965 images, validation directory - 320 images and in the testing directory - 51 images.

Data usage method:
It was decided to conduct two different experiments with the same data set but with different types of labels.
The **first experiment** is to classify images according to the presence of different types of acne on the face, without designating pixel locations. We wrote a script to change the label format and in the end, we created a csv file with a table. To understand what the table looks like, look at *"Table 1"*.
The essence of the **second experiment** is that we train the model using information about the location of each pimple on a person's face. Thus, when training, the model will see the exact location of the pimples as shown in *"Figure 2"*.

	Name	0	1	2	3
0	levle0_427_png_jpg.rf.00283cbf0ece18facfc50661...	1	0	1	0
1	levle0_514_jpg.rf.00240c05ed2ff4c06082ebace36b...	1	0	0	0
2	levle1_384_png_jpg.rf.00a624146dd38192a31ff3fd...	1	1	1	1
3	levle1_505_png_jpg.rf.00cb23b99ce6ed9efb9ba5ab...	1	0	1	1
4	levle1_295_jpg.rf.00de08138cf6586e349341a376bd...	1	0	1	0

Table 1. First five examples of image label notation

Figure 1. Sample images from the training date (Experiment I)

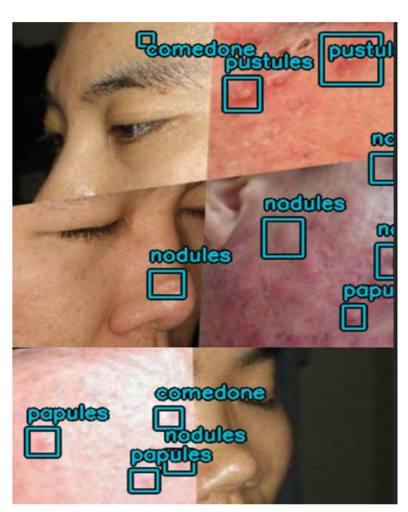


Figure 2. Sample images from the training date (Experiment II)

MODEL EXPERIMENTS

Preparing data for training:
The code loads images from the training, validation, and test directories, resizes them to a target size of (416, 416, 3), and normalizes their pixel values to the range [0, 1]. It then stores the processed images in lists X_train, X_val, and X_test. Finally, it converts these lists into NumPy arrays and prints their shapes. This preprocessing step prepares the image data for input into a machine learning or deep learning model.

Data augmentation:
The code defines a data augmentation pipeline using Keras Sequential model. The pipeline consists of several augmentation layers:
- RandomFlip
- RandomShear
- JitteredResize

Early Stopping:
Early stopping is a technique used in Deep learning to prevent overfitting during training by monitoring a metric, such as validation loss or accuracy, and stopping the training process when the performance on a validation set starts to degrade. Instead of training for a fixed number of epochs, early stopping allows the training to halt once it detects that further training may lead to worse generalization on unseen data.

Model Training:

CNN Model:
The model is a convolutional neural network (CNN) designed for identifying and classifying different types of acne lesions. It consists of several convolutional layers followed by max-pooling layers for downsampling, batch normalization layers for normalization, and dropout layers to prevent overfitting. The final layers include fully connected (dense) layers with ReLU activation functions and dropout regularization, culminating in an output layer with sigmoid activation for multi-label classification. The model is compiled using the Adam optimizer and binary cross-entropy loss function. Here's a summary of the architecture[6]:

- Input shape: (416, 416, 3)
- Convolutional layers with increasing filter sizes (16, 32, 64) and ReLU activation
- Max-pooling layers (2x2)
- Batch normalization layers
- Dropout layers with a dropout rate of 0.2 or 0.5
- Fully connected layers with ReLU activation (128, 64)
- Output layer with sigmoid activation for multi-label classification (4 classes)
- Compiled with Adam optimizer and binary cross-entropy loss

Pre-trained Models:
The models are based on different types of architecture (such as VGG19, ResNet50, EfficientNet B3), pre-trained on the ImageNet dataset. The pre-trained weights are frozen to retain the learned features. The architecture is then extended with additional layers for fine-tuning on the acne lesion classification task. A flatten layer is added to transform the output of ResNet50 into a one-dimensional vector, followed by a dropout layer to prevent overfitting. Finally, a dense layer with sigmoid activation is added for multi-label classification, predicting the presence of four types of acne lesions. The model is compiled using the RMSprop optimizer with a learning rate of 1e-4 and binary cross-entropy loss function.

YOLO Model:
The code snippet initializes a YOLOv8 object detector model. It utilizes a YOLOv8 small backbone pre-trained on the COCO dataset[4]. The detector model is configured with the number of classes in the dataset, bounding box format, backbone architecture, and FPN depth. It is compiled using the Adam optimizer with a specified learning rate and global clipnorm. The loss functions for classification and box regression are defined as binary cross-entropy and complete IoU (ciou) loss, respectively. Finally, the model is trained using the fit method with training and validation datasets for a specified number of epochs.

EXPERIMENTAL RESULTS

Experiment I results:

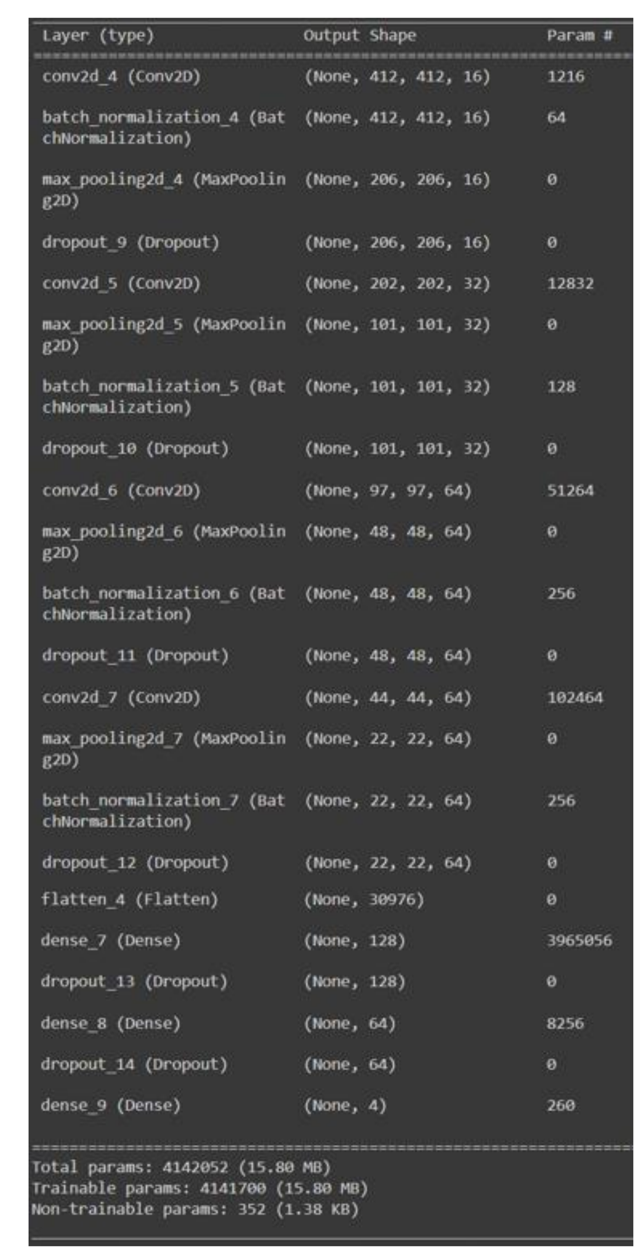
As can be understood from the results of our first experiment, the results of the first model are the best. The Accuracy of this model is about 82.35% and the loss is about 0.68. To achieve such indicators, we used the CNN model without the Pre-trained model in the first. This model is a special model for multi-label problems. To have an understanding of exactly how the model is built, you can take a look at it in "Figure 3". After we launched this model for 20 epochs, it was decided to reduce their number to do an Early Stop. As a result, it turned out that the model showed the same accuracy in the test, but the loss became slightly smaller (0.63).

In other cases, pre-trained models are used, which do not perform well under loss. From "Table 2" you can see the results of accuracy and loss obtained from the test dataset. Even though all pre-trained models are not suitable for solving a problem of this type, the VGG19 model did the best job, showing a loss result of 2.15. During the experiment, we carried out fine-tuning for each model; one of the results turned out to be no better than the previous ones.

Experiment II results:

At this stage of the code, we can conclude that with the help of information and training materials, we were able to achieve the correct construction of a model with a multilabel problem. In "Figure 5" you can see how well the trained model does. Blue squares indicate acne labels and yellow squares are the predictions of our model.

An indicator in the YOLO model is box loss, which averages 2.50, which is a good result. "box loss" is a value that shows the accuracy of finding the midpoint of an object, in our case the middle of a pimple. The "Class loss" indicator, which shows the accuracy of the classification of objects in the image, is 0.67. Despite these results, the loss for this model is 3.17. This is a satisfactory result for a multi-label classification model, but it is still an indication that the model cannot cope with the task as well as possible[3].



Layer (type)	Output Shape	Param #
conv2d_4 (Conv2D)	(None, 412, 412, 16)	1216
batch_normalization_4 (BatchNormalization)	(None, 412, 412, 16)	64
max_pooling2d_4 (MaxPooling2D)	(None, 206, 206, 16)	0
dropout_9 (Dropout)	(None, 206, 206, 16)	0
conv2d_5 (Conv2D)	(None, 202, 202, 32)	12832
max_pooling2d_5 (MaxPooling2D)	(None, 101, 101, 32)	0
batch_normalization_5 (BatchNormalization)	(None, 101, 101, 32)	128
dropout_10 (Dropout)	(None, 101, 101, 32)	0
conv2d_6 (Conv2D)	(None, 97, 97, 64)	51264
max_pooling2d_6 (MaxPooling2D)	(None, 48, 48, 64)	0
batch_normalization_6 (BatchNormalization)	(None, 48, 48, 64)	256
dropout_11 (Dropout)	(None, 48, 48, 64)	0
conv2d_7 (Conv2D)	(None, 44, 44, 64)	102464
max_pooling2d_7 (MaxPooling2D)	(None, 22, 22, 64)	0
batch_normalization_7 (BatchNormalization)	(None, 22, 22, 64)	256
dropout_12 (Dropout)	(None, 22, 22, 64)	0
flatten_4 (Flatten)	(None, 30976)	0
dense_7 (Dense)	(None, 128)	3965056
dropout_13 (Dropout)	(None, 128)	0
dense_8 (Dense)	(None, 64)	8256
dropout_14 (Dropout)	(None, 64)	0
dense_9 (Dense)	(None, 4)	260

Total params: 4142052 (15.80 MB)
Trainable params: 4141700 (15.80 MB)
Non-trainable params: 352 (1.38 KB)

Figure 3. Detail of the CNN model without pre-trained model in Experiment I

Indicators	MODEL CNN	MODEL (ResNet50)	MODEL (VGG19)	MODEL (EfficientNetB3)
Loss	0.63	36.75	2.15	14.62
Accuracy	82.35%	82.35%	68.63%	82.35%

Table 2. Models of Experiments I (Test loss/acc)

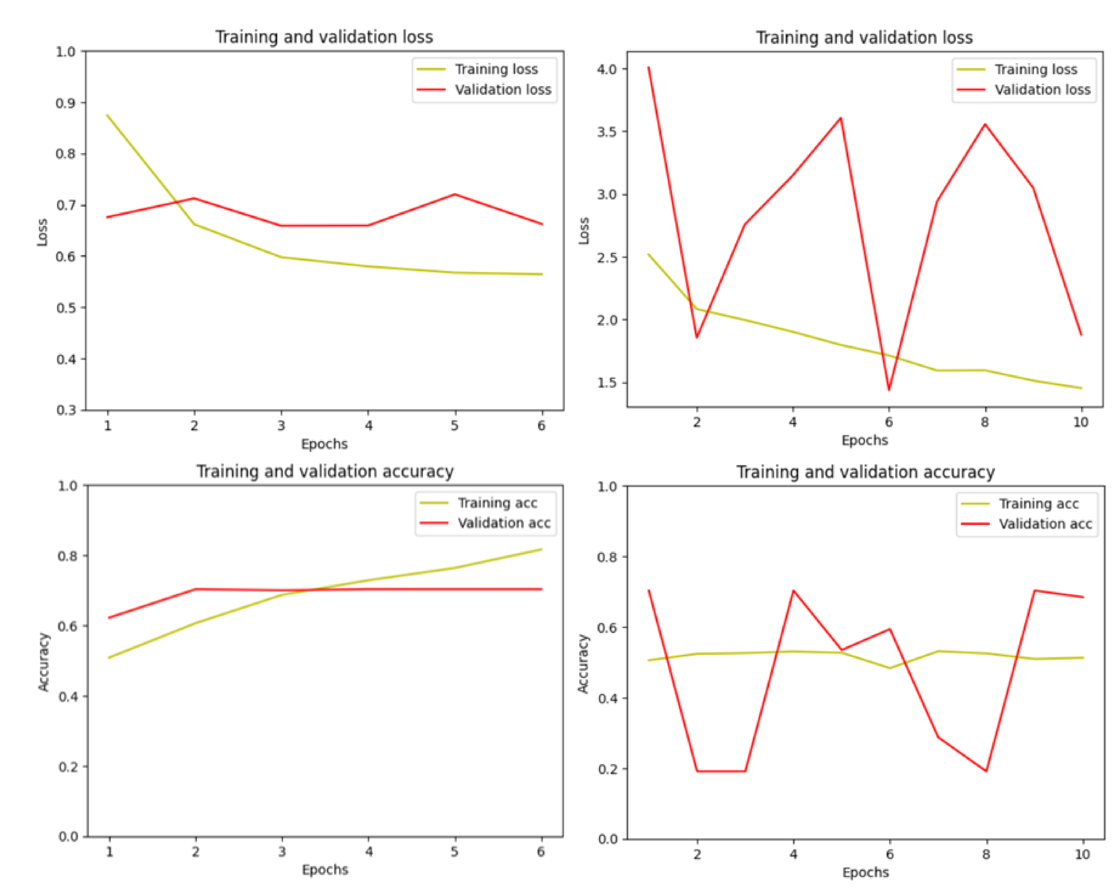


Figure 4. Graphs of Training Validation Loss/Accuracy of the CNN model and the CNN + ResNet50 pre-trained model

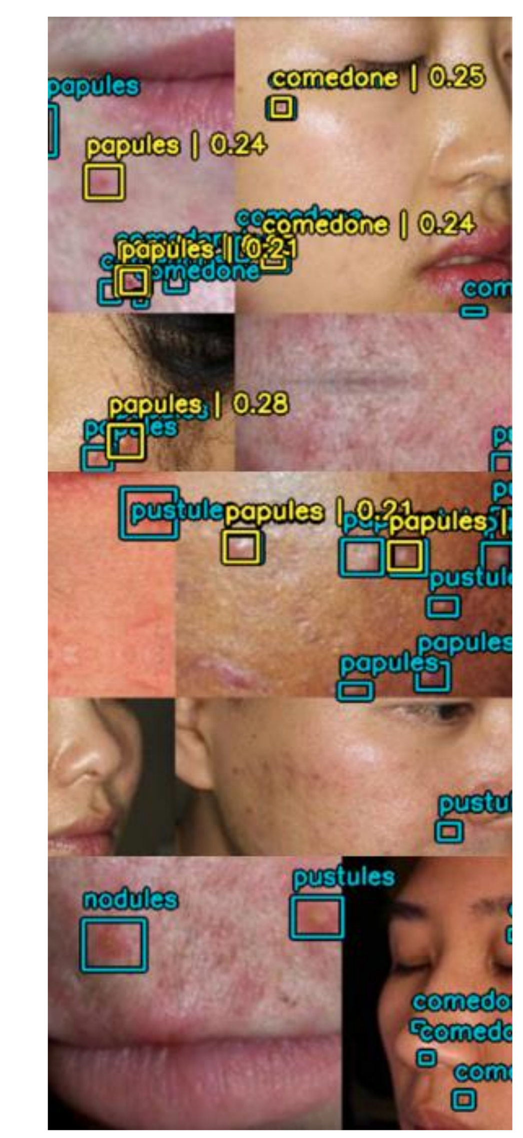


Figure 5. Prediction performance of the YOLO-trained model (Experiment II)

DISCUSSION and CONCLUSION

The CNN model without pre-trained models shows much better results:
The results of the first experiment demonstrate the effectiveness of the CNN model without pre-trained models for acne lesion classification. With an accuracy of approximately 82.35% and a lower loss compared to pre-trained models, this approach proves its suitability for the task. We think that the reason that pre-trained models are not suitable for our task is the fact that these models were trained to classify objects that do not look like pimples

The first experiment shows better results than the model from the second experiment:
Experiment I, utilizing the CNN model without pre-trained models, outperformed Experiment II, which employed YOLO models. Despite YOLO's capabilities in object detection, the CNN model exhibited superior performance in the multi-label classification of acne lesions. However, it should be taken into account that in the first experiment we will only have the presence of certain acne on the face, while the task of the second experiment is to indicate the location of each pimple and classify it.

Lack of other datasets with acne classification:
One limitation encountered in this research is the lack of diverse datasets specifically annotated for acne lesion classification. This limitation may have influenced the model's performance and generalizability. Future efforts should focus on curating and annotating additional datasets to address this gap.

The training performance is not high, since the GPUs that were used in this research work often failed to cope with the tasks and failed:
In the project were used Google Colab GPUs (A100 GPU, V100 GPU, L4 GPU, T4 GPU). The inadequate training performance attributed to GPU limitations, particularly in Google Colab, highlights the importance of adequate computational resources in deep learning research. Access to more powerful GPUs or alternative computing platforms could potentially alleviate these challenges and improve training efficiency.

YOLO models have a hard time recognizing small objects like pimples:
The difficulty of YOLO models in recognizing small objects like pimples highlights a common challenge in object detection algorithms. YOLO's architecture, which divides the input image into a grid and predicts bounding boxes and class probabilities for each grid cell, may struggle with small objects due to their limited spatial extent within the grid cells. This limitation can lead to decreased accuracy and localization precision for small objects, such as pimples on the face.
To address this challenge, further research and experimentation are necessary to optimize YOLO models or explore alternative approaches tailored to the specific characteristics of acne lesions.

It's worth delving deeper into computer vision technologies and integrating work into Roboflow:
Integrating the research into platforms like Roboflow offers opportunities for collaboration, dataset management, and further exploration of computer vision technologies. Leveraging such platforms can facilitate access to resources, tools, and expertise, ultimately enhancing the research outcomes and impact.

In conclusion, while the CNN model without pre-trained models demonstrates promising results for acne lesion classification, there are still challenges and opportunities for improvement, including dataset diversity, computational resource limitations, algorithm optimization, and integration with specialized platforms. Addressing these aspects can contribute to advancing research in Deep Learning and Computer Vision for dermatological applications.

FUTURE WORK

Acne recognition technology has promise, especially in the context of automating and improving acne diagnosis and treatment:

Personalized Treatment Plans: By accurately classifying different types of acne lesions, the technology can help dermatologists tailor treatment plans to each patient's specific needs. For example, a patient with predominantly nodular acne may require a different treatment approach than someone with primarily papules and pustules. Personalized treatment plans can improve efficacy and reduce the risk of scarring.

Remote Monitoring and Telemedicine: Remote monitoring can improve accessibility to healthcare, particularly for individuals in rural or underserved areas.

Integration with Skincare Devices: As technology continues to converge, we may see the integration of deep learning algorithms with skincare devices such as smartphone apps or handheld scanners. These devices could provide real-time feedback on acne severity, recommend skincare products, or even deliver targeted treatments like light therapy. Such innovations could empower individuals to take proactive steps in managing their acne.

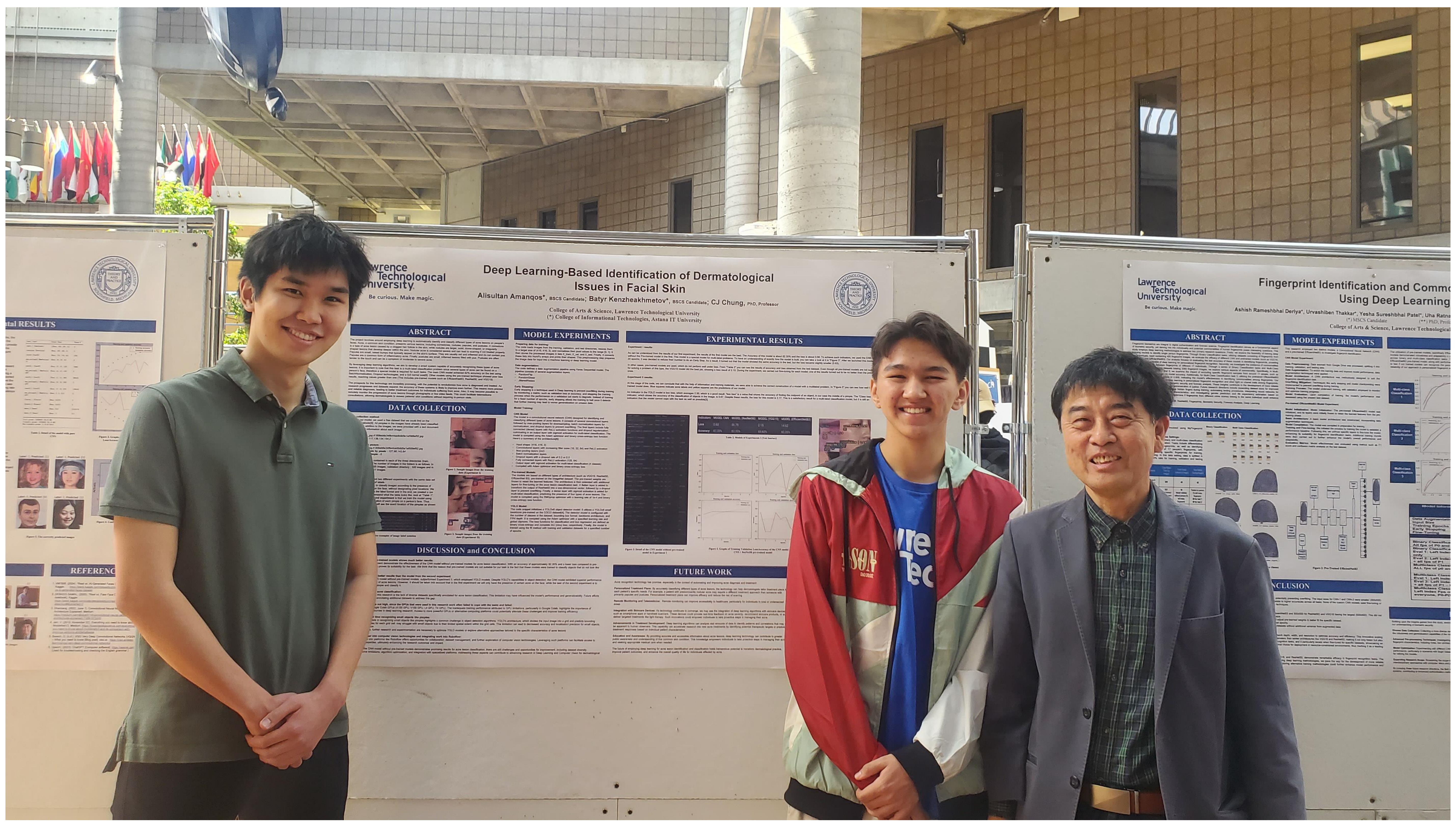
Advancements in Treatment Development: Deep learning algorithms can analyze vast amounts of data to identify patterns and correlations that may not be apparent to human observers. This capability can accelerate research into new acne treatments by identifying potential therapeutic targets or predicting treatment responses based on individual patient characteristics.

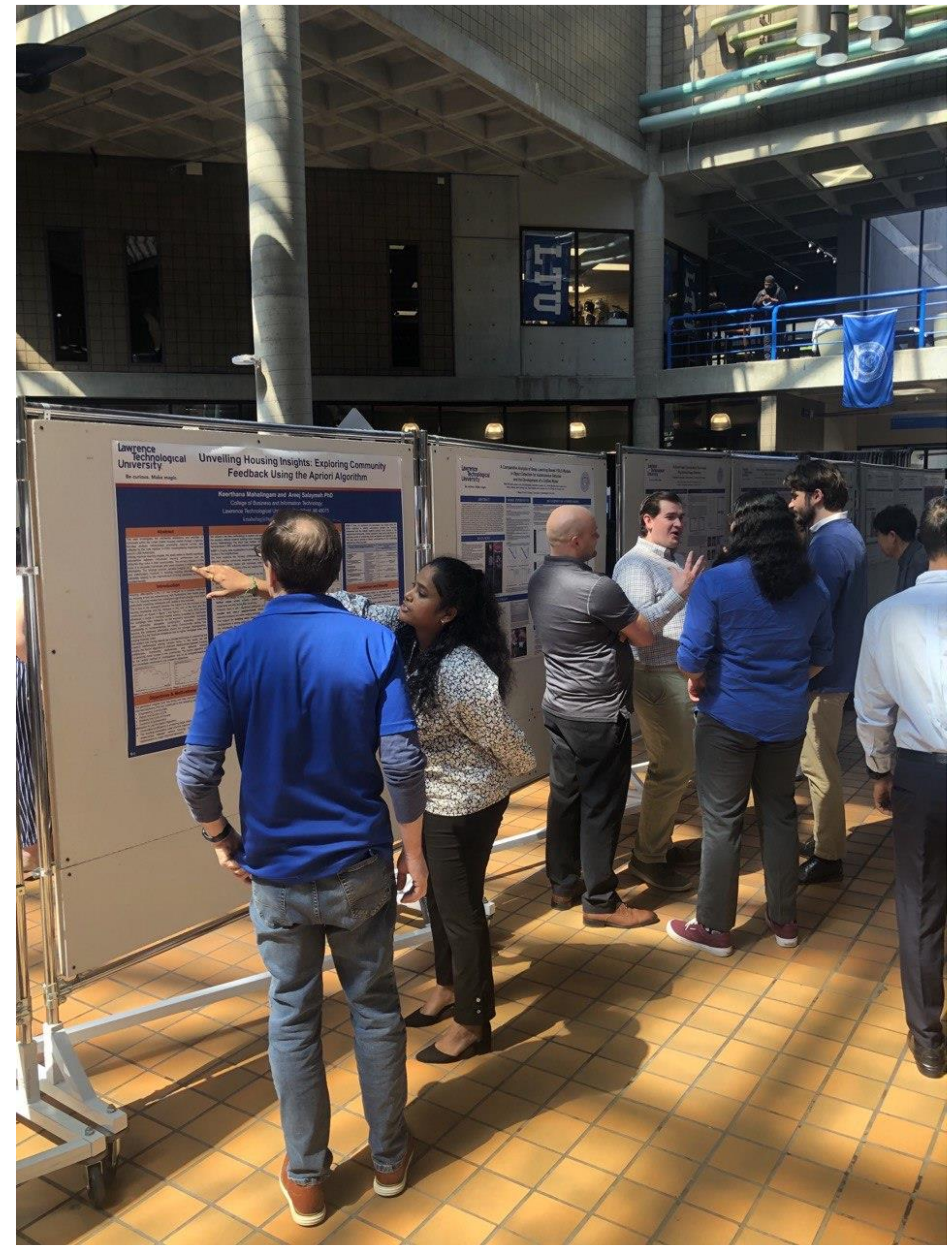
Education and Awareness: By providing accurate and accessible information about acne lesions, deep learning technology can contribute to greater public awareness and understanding of this common skin condition. This knowledge empowers individuals to take proactive steps in managing their acne and seeking appropriate medical care when needed.

The future of employing deep learning for acne lesion identification and classification holds tremendous potential to transform dermatological practice, improve patient outcomes, and enhance the overall quality of life for individuals affected by acne.

REFERENCES

1. Lakewood, I. Stenbit, T. Patel. (2023/04/08). *"Object Detection with KerasCV".* https://keras.io/guides/keras_cv/object_detection_keras_cv/#model-creation
2. G. Chawda. (2023/06/26). *"Efficient Object Detection with YOLOV8 and KerasCV".* https://keras.io/examples/vision/yolov8/
3. S. Rath. (2023/09/26).*"Object Detection using KerasCV YOLOv8".* https://learnopencv.com/object-detection-using-kerascv-yolov8/
4. *"Deep Learning in Medical Imaging: Multi-label Classification with PyTorch".* https://www.youtube.com/watch?v=YsytDf-MDN4
5. *"Multilabel classification using Keras*". https://www.youtube.com/watch?v=hraKTseOuJA&t=933s
6. *"Multiclass classification using Keras".* https://www.youtube.com/watch?v=obOjpVdO3gY
7. Dataset source: https://universe.roboflow.com/acnedet/acnedet-v1/dataset/2


Lawrence Technological University
Deep Learning-Based Identification of Dermatological Issues in Facial Skin
ABSTRACT
MODEL EXPERIMENTS
EXPERIMENTAL RESULTS
DATA COLLECTION
DISCUSSION and CONCLUSION
FUTURE WORK
Lawrence Technological University
ABSTRACT
MODEL EXPERIMENTS
DATA COLLECTION


Lawrence Technological University
Unveiling Housing Insights: Exploring Community Feedback Using the Apriori Algorithm