Skin Cancer Classification using Delaunay Triangulation and Graph Convolutional Network

Caroline Angelina Sunarya[1], Jocelyn Verna Siswanto[2], Grace Shirley Cam[3], Felix Indra Kurniadi[4]
Data Science Program-School of Computer Science, Bina Nusantara University, Jakarta, Indonesia, 11530[1, 2, 3]
Computer Science Department-School of Computer Science, Bina Nusantara University, Jakarta, Indonesia, 11530[4]

***Abstract*—Oftentimes, many people or even medical workers misdiagnose skin cancer, which may lead to malpractice and thus, resulting in delayed recovery or life-threatening complications. In this research, a Graph Convolutional Network (GCN) method is proposed as a classification model and Delaunay triangulation as its feature extraction method to classify various types of skin cancers. Delaunay triangulation serves the purpose of boundary extraction, and this implementation allows the model to focus only on the cancerous lesion and ignore the skin around it. This way, the types of skin cancer can be predicted more accurately. Furthermore, GCN offers many advantages in medical image analysis over traditional CNN models. GCN can model interactions between different regions and structures in an image and perform messaging between nodes, whereas CNN is not explicitly designed to do such thing. Other than that, GCN can also leverage transfer learning and few-shot learning techniques to address the challenges of limited annotated medical image datasets. However, the result shows that the proposed model tends to overfit and is unable to generate correct predictions for new skin cancer images. There are several reasons that could lead the model to overfit, such as imbalance data, incorrect feature extraction, insufficient features for data prediction, or the data containing noise.**

Keywords*—Skin cancer; Delaunay triangulation; graph convolutional network; GCN; multilabel image classification; convolutional neural network; CNN

I. Introduction

Skin cancer is the most prevalent form of cancer for people with white or light-colored skin all over the world. Exposure to ultraviolet radiation (UVR) is a primary etiologic agent in the emergence of skin cancer, which may result in DNA damage and genetic abnormalities, subsequently leading to skin cancer [1]. In general, skin cancers are classified as melanoma or non-melanoma. Skin cancer, which includes both melanoma and non-melanoma, are the particularly prevalent form of cancer malignancies in the white population. However, the one responsible for the majority of deaths caused by cancer is melanoma skin cancer [2].

Melanoma screening is not suggested as a population screening tool for wide range of reasons. This is because the effectiveness of early detection or screening programs has not yet been tested in randomized trials [3]. Although the majority of at-risk patients attend medical appointments or even a doctor's appointment that offers the option of skin testing in melanoma detection, only a few medical practitioners have specialized knowledge or training in melanoma detection. In the study published by Annals of Oncology, researches compared the performance of 58 international dermatologists and found CNN's skin cancer diagnoses to be more accurate than diagnoses made by a panel of dermatologists. [4].

Considering that not many non-dermatologist healthcare workers and dermatologists could correctly identify melanomas, it could harm patients with skin cancer. Not knowing that one has skin cancer, patients would carry on with their lives and consider it as a normal mole, little that they know that it is a skin cancer. Ignoring the first signs of skin cancer would let the cancer grow malignant, and eventually will harm the patients. Even when healthcare workers successfully diagnose a patient with skin cancer, there are still possibilities of human error such as misdiagnosing skin cancer type (melanoma or non-melanoma, malignant or benign) which could further harm the patients through medical malpractice, namely giving the inappropriate treatment which would result in delayed recovery, and negative medical outcomes (additional pains or even other non-life threatening and life-threatening complications).

Misdiagnosis and malpractice only harm patients in a way that is very dangerous for that person and even would cost a person's life. Considering that these issues happened in real life and could even happen to ourselves, this study strives to help medical workers and patients overcome these issues. With the help of current technology, it will aid medical personnel in identifying both malignant and benign melanoma skin cancers, as well as skin cancers that are not melanoma related. Time and effort are required to gain experiences. Medical professionals have spent years researching and treating patients in a manner that only they can. Several researches have proven that modern technology outperforms medical professionals in terms of knowledge. In some cases, if not all, a doctor's opinion is needed rather than relying solely on technology. Additionally, technology can be used as a tool to substantiate medical opinion with expert advice [5].

Several research projects on the automated classification of melanoma imagery utilizing computer vision and machine learning algorithms have been carried out. Despite the fact that these researches show promising outcomes, the application of computer vision and traditional machine learning has a significant impact on classification performance through features identified in skin lesion segmentation results and classification methods [5]. This study proposes the fusion of Delaunay triangulation and graph convolutional network (GCN) as a method to classify images of skin cancer types.

The use of deep learning hoped to reduce human intervention and improve accuracy.

Both methods used were hoped to give better results in classifying skin cancer. Delaunay triangulation is used to segment cancer from the skin or to extract masks from lesion regions without requiring a training phase so that the neural network can better classify the type of skin cancer. Whilst convolutional neural network (CNN) has emerged as the most advanced network for pattern identification in medical image analysis, it is not without limitations [6], graph convolutional network (GCN) can handle irregular and non-grid data. Many real-world applications, such as social analytics, identifying fraudulent activity, traffic forecasts, computer vision, and others, generate graphs naturally. Data can be shown as graphs that encode structural details to represent relationships between items and provide more interesting insights underlying the data [7].

The data used in this study are skin cancer images obtained from the Kaggle website. The images are then resized and go through further data preprocessing steps. Skin cancer mask points were then generated using the Delaunay triangulation method. The obtained points were stored in an array and used in the modeling step using the GCN model. Delaunay triangulation serves the purpose of boundary extraction, and this implementation allows the model to focus only on the cancerous lesion and ignore the skin around it. Furthermore, GCN offers many advantages in medical image analysis over using traditional CNN models. GCN can model interactions between different regions and structures in an image and perform messaging between nodes. Other than that, GCN also can leverage transfer learning and few shot learning techniques to address the challenges of limited annotated medical image datasets.

By using cancer lesion points obtained using Delaunay triangulation and GCN as its model, classification tasks may be performed more accurately, even when dealing with images with significant noise that can lead to overfitting. When using CNN as the model, the images were processed as it is without any further data preprocessing steps. CNN itself is considered as the most advanced network for pattern identification in medical image analysis. However, when faced with images of skin cancer taken from quite a distance, the model would not perform well. Rather than focusing on the cancer lesion alone, the model would also focus on the skin surrounding it. The model would attempt too hard in seeking the most suitable fit to the data which will lead to overfit, as not everyone has the same skin color.

II. Related Work

There are many studies relevant to the proposed topic of this research. One example is a study by Shi Yin and friends who performed the diagnosis of kidney disease using ultrasound imaging [8]. In this study, the researchers used a methodology that builds on recent advances in Deep MIL, employing a convolutional neural network (CNN) to extract instance-level data from 2D kidney ultrasound images and a graph convolutional network (GCN) examine the characteristics of and further optimize the instance level. It works by looking at possible correlations between instances of the same bag. ReLU activation function was used in this study. GCNs, knowledge-based MIL collection, and instance-level monitoring based on instances with reliable labels can improve MIL classification performance. The dataset used in this study was based on a dataset of US clinical renal scans from renal patients collected at the Children's Hospital of Philadelphia (CHOP). This study shows that their proposed method obtained an accuracy of 84.89%.

Other research on skin cancer detection is proposed by Abilash Panja, Christy Jackson J, and Abdul Quadir [9]. The researchers presented a deep convolutional neural network (CNN) to classify melanoma images into either benign or malignant groups. The model is constructed with convolution layer, activation layer, batch normalization layer, and pooling layer, with each of the convolutional layers has an activation function that uses ReLU, while the output layer has an activation function that is sigmoid, and the use of binary cross entropy for model fitting. Deep neural network's stability and performance are improved by using a normalization layer. The dataset used in this work originates from the ISIC archive repository dataset, which includes 2637 input photos for training from two benign and harmful classes, as well as 800 test images. The research compared the epoch and loss function performance differences with 100 random images, and has achieved 83.38% accuracy with 10 epochs, 87.52% with 25 epoch, 91.21% with 50 epoch, and 95.61% with 100 epoch. A dropout value of 0.5% was used as a constant in all experiments, giving the best performance at 100 epochs with 95% accuracy.

Research by Muhammad Arif and friends also predicts skin cancer using two different datasets: the datasets associated with the Interactive Atlas of Dermoscopy (EDRA) and the newly suggested International Skin Imaging Collaboration (ISIC), 2017. There are 2000 photos accessible for training, with only about 400 of those identified as melanoma. It also includes 150 images that can be utilized as the test set and 600 images to function as the set used for validation. This research uses ADF techniques in image processing to reduce noise, K-means clustering, and modified K-means clustering. In this research, modified K-means clustering shows better results than other methods with accuracy up to 0.9992%. This study discovered the fact that the neural network hybrid CNN (CNN+IHHO) classifier outperforms other approaches that includes ResNet, VGG, and DenseNet, with a performance rating of 97.3%. [10].

In most researches, the Convolutional Neural Network is the most common method to use when classifying images and can give more than 80% accuracy. However, this research proposes a fusion of Delaunay Triangulation method and Graph Convolutional Network (GCN) to predict skin cancer. Delaunay Triangulation is used to extract the cancer lesion areas and separate it from the healthy skin area. The extracted feature is used to train the GCN model, in hopes that the model can provide higher accuracy than previous studies by using only the cancer lesion area.

III. METHODOLOGY

A. Dataset Description

The data was collected by International Skin Imaging Collaboration (ISIC) through Kaggle platform [11], consisting of 2357 skin cancer images of over 2000 patients from various sources such as Hospital Clínic de Barcelona, Medical University of Vienna, Memorial Sloan Kettering Cancer Center, Melanoma Institute Australia, The University of Queensland, and the University of Athens Medical School in 2020 [12]. There are two folders in the dataset, Training and Testing. Training folder is composed of nine folders with the total of 2239 images, representing nine types of skin cancer disease: Actinic Keratosis, Basal Cell Carcinoma, Dermatofibroma, Melanoma, Nevus, Pigmented Benign Keratosis, Seborrheic Keratosis, Squamous Cell Carcinoma, and Vascular Lesion. Testing folder also consists of nine folders with a total of 118 images. However, this research only uses the first five folders and classifies those five types of skin cancer disease due to insufficient resources of RAM and memory.

B. Proposed Methodology

The proposed methodology consists of (1) Exploratory Data Analysis, (2) Data Preprocessing, and (3) Modeling and Evaluation. The flow of the process is shown in Fig. 1.

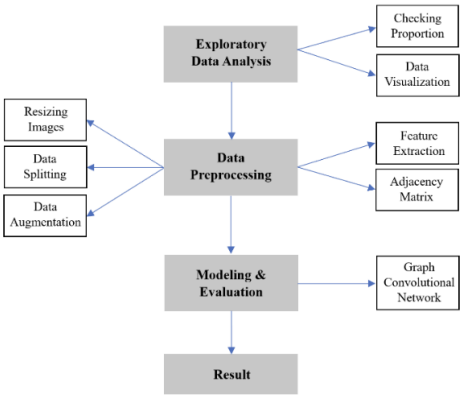

Fig. 1. Proposed methodology.

1) Exploratory data analysis: Research begins with importing and loading the data into the software, and then randomly visualizing nine images to get an overall depiction of the dataset. The software used for this research is Google Colab, with Python as the programming language. The process of visualizing the dataset and retrieving insights from it was done using the functions and libraries that are available in Python such as Matplotlib and Tensorflow Keras. After the data visualization process, the proportion of each skin cancer class is checked to decide whether the data needs to be balanced through data augmentation process or left as is.

2) Data preprocessing: The process consists of resizing all images into 180x180 pixels dimension, data splitting, data augmentation for training set, feature extraction process, and adjacency matrix creation. In the data splitting process, data is divided into three sets, namely the training set, validation set, and testing set. Training set will be used for modeling process and validation set will be used for model evaluation. After splitting, the data augmentation process is carried out to add the number of images in each class to balance the training set. Then, a feature extraction process is carried out using Delaunay Triangulation method to generate the graph representation of each cancer image along with the average pixel values inside the graph, which can help differentiate between healthy skin and cancer skin region. Lastly, adjacency matrices are created based on the Delaunay points to describe the connectivity between nodes in each graph.

3) Modeling and evaluation: A Graph Convolutional Network model was created and trained using the pixels and adjacency matrix of all images in the training set as inputs. After that, the model is evaluated using the validation set. During evaluation process, the model is being analyzed whether it was overfitting or underfitting. The model result was compared with the previous result and was then modified a few times to improve its performance.

4) Result: After finding a model with the best performance, the model was then implemented to the testing set to predict the outcomes.

C. Delaunay Triangulation

Delaunay triangulation is a method of dividing a region into sub regions in triangular shapes [13]. This method generates a graph representation of the cancer shape in each image, which can be used as an input in Graph Convolutional Network. The graph will make it easier to distinguish the cancer skin from the normal skin as well as measuring the size of the cancer region.

The Delaunay triangulation implemented in this research is based on Quickhull Algorithm to compute the 2D convex hull of a set of points that are generated from the contour of each image, with the process as follows:

- Each image is converted into Grayscale color to find its contour.
- A set of points (100 points) are extracted from the contour of each image.
- Convex hulls are computed from those points using Quickhull Algorithm below [14]:

a) Find maximum and minimum points along the x dimension of that set of points.

b) Add the minimum and maximum points to the convex hull.

c) Connect the maximum and minimum points into a line.

d) Divide the remaining points into two subsets based on whether they are above or below the line.

e) Find points that are the farthest from the line and add them to the convex hull.

f) Construct a triangle using the line and the farthest point.

g) Remove all points in the subset that lies inside the triangle.

h) Step d to g are carried out recursively until no points remain in any subset.

i) The remaining set of points in the convex hull represents Delaunay triangulation.

- The average pixel values of each Delaunay triangle are extracted.
- Lastly, adjacency matrices are created based on the Delaunay points. Adjacency matrices have rows and columns labeled by graph vertices, with the number 0 or 1 in position (xi , xj) based on whether xi and xj are adjacent (1) or not (2), respectively [15].

D. Graph Convolutional Network

In this research, Graph Convolutional Network is used as a model to predict the types of skin cancer. Graph Convolutional Network (GCN) is a type of multilayer neural network that directly operates on graph-structured data, where it produces a vector representation of each node based on the features of its surrounding node [16]. GCN is used as it is suitable for inputs generated by Delaunay Triangulation, which is in the form of graphs (vertices, edge, and adjacency information for the triangles).

In the research, the GCN model propagates values forward through five layers using the Forward Propagation equation. The equation can be seen below (1):

$$H^{[i+1]} = \alpha(W^{[i]} H^{[i]} \bar{A}) \qquad (1)$$

where $H^{[i+1]}$ is the feature representation of layer i+1, is the activation function applied for the inputs in each layer, $W^{[i]}$ is the weight values for layer i, $H^{[i]}$ is the feature representation of layer i, and $\bar{A}$ is the normalized adjacency matrix [17].

IV. Result and Discussion

The Training dataset consists of 2239 skin cancer images distributed in nine folders. However, by using only five classes due to insufficient resources, there are only 1380 images left to use. The overall depiction of skin cancer provided by the dataset can be seen in Fig. 2.

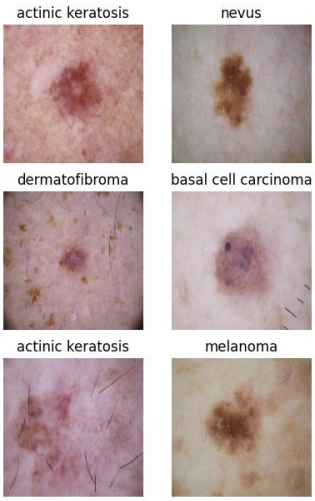

Fig. 2. Data depiction.

In the above figure, six images of skin cancer belonging to five classes are visualized using TensorFlow Keras library in Python to process the images, and Matplotlib library to visualize the processed images. The five classes of skin cancer in the dataset are Actinic Keratosis, Basal Cell Carcinoma, Dermatofibroma, Melanoma, and Nevus.

1) Artinic keratosis: Also known as Solar Keratosis, Actinic Keratosis is a precancerous skin condition caused by a long-term exposure of Ultraviolet radiation from the sun. This condition is mostly found at places that are commonly exposed to sun, such as the backs of the hands or the face, often affecting forehead, temple, balding scalp, upper or vermilion of lower lip, cheeks, ears, and nose [18].

2) Basal cell carcinoma: Basal Cell Carcinoma is a skin cancer forming in the basal cell of a skin, which is located in the lower part of epidermis (outside layer of a skin). This type of cancer has a form of shiny bump or scaly flat patch on the skin and can gradually grow over time. This condition is caused by DNA mutation, inherited gene defects, or excessive exposure of ultraviolet (UV) rays. It is most commonly found among elderly male, as well as people with fair skin, blue eyes, and blond or red hair [19].

3) Dermatofibroma: Also known as Cutaneous Fibrous Histiocytoma, Dermatofibroma is a benign fibrous nodule that appears commonly on the skin of lower legs, upper backs, and arms. This condition is mostly found among females more than males and occurs in adults of any ethnicity. Some causes of dermatofibromas are a reaction to trauma, such as insect bite or small cuts in the area where the nodule later formed [20].

4) Melanoma: Melanoma is a type of skin cancer that is produced by malignant transformations of melanocytes, which are the cells that produce pigment on the skin [21]. It is caused mainly by overexposure of Ultraviolet radiation or artificial sources like solarium, can spread to many parts of the body, and can be incurable. This cancer is mostly found in people with pale skin, moles in skin, many sunburns, or old age [22].

5) Nevus: Also called a mole, Nevus (plural: Nevi) is a patch of skin formed due to an overgrowth of cells in the epidermis (outermost layer) part of skin. Nevus is harmless and common; it can be seen at birth or develop in early childhood. Any individual, of all ages and ethnicities can develop a nevus on their skin, and as they get older, nevi can become darker and thicker or even grow into wart-like form [23].

The proportion and number of images in each class from the dataset can be seen in Table I.

TABLE I. Data Proportion

Class	*Number of Images*	*Proportion*
Melanoma	438	32%
Basal Cell Carcinoma	376	27%
Nevus	357	26%
Actinic Keratosis	114	8%
Dermatofibroma	95	7%

Result shows that the proportion of data in each class is imbalanced. Therefore, a data augmentation process is necessary to balance the data and ensure a good performance for the model. An imbalanced data can cause bias towards the model and cause lower accuracy to minority classes. Data augmentation will increase the size of those minority classes to obtain the same size as the largest class [24]. Several transformation techniques such as vertical and horizontal flipping, shifting, zooming, as well as rotating images are used when duplicating the data to increase its variety. The transformation process is carried out using ImageDataGenerator() function from Tensorflow Keras library.

After the Exploratory Data Analysis process, Data Preprocessing is carried out starting with resizing the height and width of all images into 180 pixels. Resizing images is important as images captured by cameras can vary in size, but a neural network only receives 1 input size. Thus, establishing a base size for all images is crucial so the images can be fed into the algorithms [25].

In Data Splitting process, the resized images in the Training folder are divided into Training set and Validation set, with the proportion of 80% and 20% respectively. An empirical analysis has proven that the best results are obtained by allocating 20-30% of the original dataset for validation set and the remaining 70-80% proportion for training set, as it can provide adequate data for training a model as well as avoid overfitting [26].

After splitting the dataset, Feature Extraction process is carried out using Delaunay triangulation method to generate a graph representation of each image and extract its pixels, as explained in the methodology. Delaunay triangulation method is carried out using Delaunay() function from Scipy library in Python, and its result can be seen in Fig. 3.

In the figure, the original image is shown on the left side, whereas images on the right side are the result after Delaunay triangles are generated. The Delaunay method is able to generate triangles properly in the cancer area of four classes. However, triangles in Nevus class are generated in the background instead of the area of interest. This is the weakness found when using Delaunay Triangulation method for skin cancer data.

In this research, adjacency matrices are created using NumPy library in Python, with the adjacency information based on the graph formed by Delaunay triangles. The distance of all points on the graph are calculated using Euclidean Distance formula, and two points are considered connected or adjacent if the distance is greater than 2.5 pixels. The Euclidean Distance method uses Pythagorean theorem [27], with the equation as seen below (2):

$$d_{ij} = \sqrt{(x_i - x_j)^2 - (y_i - y_j)^2} \quad (2)$$

where dij is the distance between position i and j, (xi ,yi) is the coordinate of a point in position i, and (xj ,yj) is the coordinate of a point in position j.

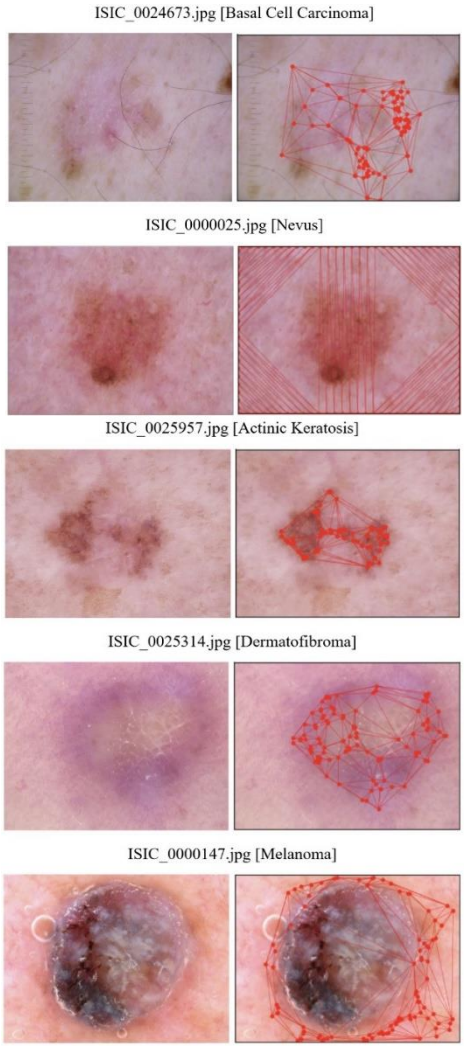

Fig. 3. Delaunay points.

After generating the graph representation of each image, a GCN model is built to receive two inputs, which are the adjacency matrix and pixels matrix of each image. The model is constructed using 5 Graph Convolutional layers, a batch normalization after each Graph Convolutional layer, 1 flatten layer, and 1 output layer. Each graph convolutional layer has 32, 64, 128, 256, and 512 filters respectively, and all layers use ReLU activation function as well as L2 regularization method.

Batch normalization layers are added as they can normalize the output of the layer. Input that has been normalized can increase stability of the optimization process, which helps improve model's performance [28]. L2 regularization is added in each Graph Convolutional layer as it can do regularization to help overcome overfitting problem. L2 regularization can learn complex patterns well, so it is suitable for data in the form of images, which have complex patterns [29].

A ReLU activation function is used in all Graph Convolutional layers as it is the most often used function in the neural network model and can overcome the vanishing gradient problem, which occurs when the gradients of activation function become very small during backpropagation process and causes very small updates in the weights of earlier layers,

so the model learns slowly or not learning at all. This problem may occur when using sigmoid and hyperbolic tan (tanh) activation functions [30]. The graph of ReLU function can be seen in Fig. 4 [31], and the formula of ReLU is defined below (3):

$$f(x) = \max(0, x) \quad (3)$$

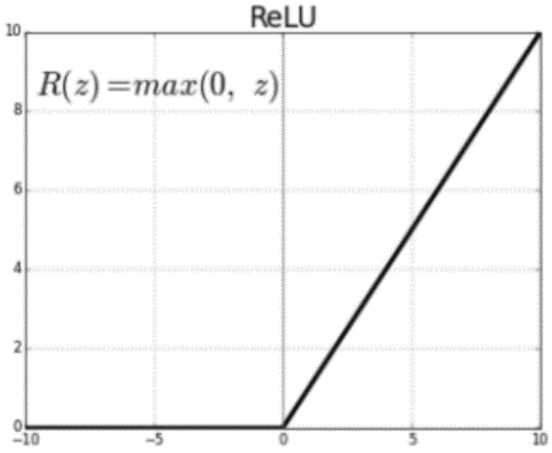

Fig. 4. ReLU activation function.

The model is compiled using 'Adam' optimizer and 'categorical_crossentropy' as its loss function. Categorical cross entropy is specifically used for multi-class classification problems by taking class 1, 2, or n shaped labels, whereas the Adam optimizer is used because for classification problems, it often performs better at generalizing than other optimizers [32]. Thus, it is better at preventing overfitting for most cases.

After compilation, the model is trained with 50 epochs and a batch size of 64, so the model can train for up to 50 iterations and utilize 64 training examples in one iteration. Once the model is trained, evaluation process is carried on. The accuracy and loss during the last epoch can be seen in Table II.

TABLE II. ACCURACY AND LOSS OF LAST EPOCH

	Accuracy	***Loss***
Training Set	0.66	1.22
Validation Set	0.32	3.10

The accuracies and losses of training and validation set during the model training using all epochs from 1 to 50 can be seen in Fig. 5 and 6. During the training process, the accuracy of the model significantly improved, and the loss values were stable throughout each epoch. However, the accuracy of the validation process increased or decreased significantly during the modeling process. The loss value of the validation process decreased significantly from epoch 0 to approximately epoch 1 and experienced a fairly large increase and decrease from approximately epoch 1 to epoch 7.

The above Accuracy and Loss Graph is based on the range of epochs and visualized using Matplotlib library in Python. In the graph, it can be seen that the accuracy for training set kept increasing until it reached more than 60%, showing that the model works quite well in predicting the training set. However, the accuracy difference between training and validation set gradually increases, as the accuracy of validation set makes no significant improvement during each epoch. Thus, this shows that the GCN model tends to overfit, even though the loss value during training process of each epoch kept decreasing.

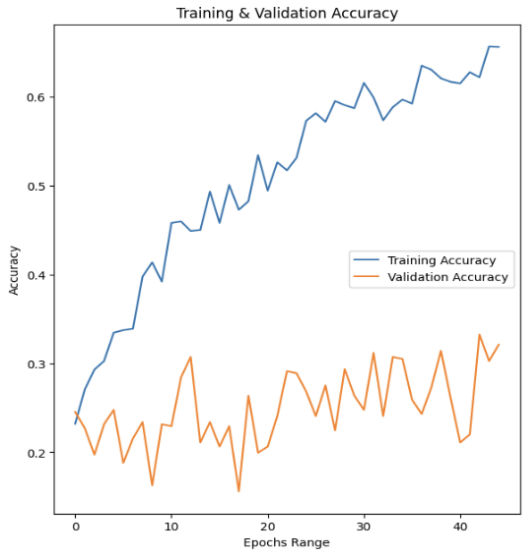

Fig. 5. Training & validation accuracy.

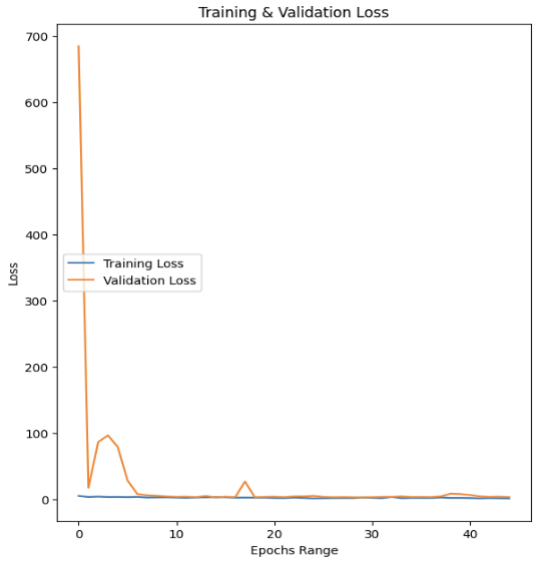

Fig. 6. Training and validation loss.

After evaluating the model's performance, the model is used to predict the testing set. Some prediction results can be seen in Table III.

The last accuracy result for the validation process is around 32%, which causes unreliable prediction results for the testing set. From the table, it can be seen that the image with the name ISIC_0024411 is predicted to be in Dermatofibroma. In reality, the image belongs to Basal Cell Carcinoma class. Moreover, the image with the name ISIC_0024403 and ISIC_0024431 is predicted to be Nevus when in reality, the image belongs to Basal Cell Carcinoma as well.

TABLE III. ACCURACY AND LOSS OF LAST EPOCH

No.	*ActinicKera-tosis*	*Basal Cell Carci-noma*	*Derma-tofibr-oma*	*Mela-noma*	*Nevus*	*Image name*	*Highest Proba-bility*
1	0.27	0.05	0.41	0.10	0.17	ISIC_0024411	Derma-tofibr-oma
2	0.21	0.58	0.18	0.00	0.03	ISIC_0024472	Basal Cell Carcin-oma
3	0.00	0.02	0.01	0.00	0.97	ISIC_0024403	Nevus
4	0.32	0.26	0.03	0.10	0.29	ISIC_0024454	Actinic Keratо-sis
5	0.02	0.00	0.37	0.01	0.60	ISIC_0024431	Nevus

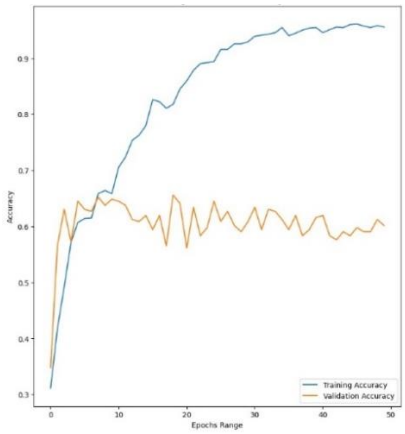

Fig. 7. Training and validation accuracy of the CNN model.

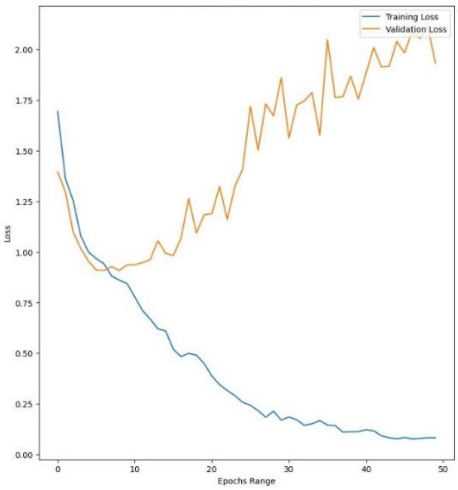

Fig. 8. Training and validation loss of the CNN model.

TABLE IV. CNN VS GCN

	CNN	*GCN*
Training Accuracy	96.55%	66%
Validation Accuracy	60.14%	32%
Training Loss	0.07	1.22
Validation Loss	1.93	3.10

When compared with the proposed model, convolutional neural network (CNN) performed better. A comparison of accuracy using the CNN model is shown in Fig. 7 and Table IV. The accuracy of the training process was 96.55% and the accuracy of the validation process was 60.14%. As shown in Fig. 8, the loss in the training process steadily decreased from epoch to epoch and did not change significantly from epoch 40 to epoch 50. However, the loss in the validation process was the other way around. From epoch 0 to approximately epoch 5, the loss of the validation process decreased significantly, and from there until the last epoch the loss increased significantly. A comparison of accuracy and loss in Table IV shows the better performance of CNN compared to the proposed model. The accuracy of the CNN model both in the training and validation process outperformed the proposed model by 30%. The gap between the training and validation processes' loss values is similar for both CNN and GCN models. However, the loss value of the proposed model is higher.

Even though it produced better accuracies and losses than the proposed model, the CNN model overfitted. Overfitting is a scenario that occurs when the predictive model fails to generalize the observed data properly in order to fit both training and testing data well. An overfitting condition takes place when the model attempts too hard in seeking the most suitable fit to the data it was trained on and adapts for noise in the data by retaining multiple training data characteristics rather than discovering a general prediction rule [33]. For CNN model, the accuracies and losses of the training process kept improving through each epoch or iteration, while the accuracy and loss of the validation process stopped improving after going through a certain amount of iteration.

V. CONCLUSION

The accuracy and loss values of CNN model are higher than the GCN model. Images used in this study were taken at close range, where cancer lesions are clearly visible, and the cancer-to-skin area ratio is nearly balanced. However, the model using CNN is overfitted. The model did not seek to fit the data well by focusing only on cancerous areas, but also on normal skin areas.

When compared with CNN, the accuracy of the proposed model is terrible, with only 66% for the training process, and 32% for the validation process. On the other hand, the accuracy of training process generated from the CNN model is 96.55% and 60.14% for validation process. Furthermore, the last loss value from the training process is lower than the validation process. As the model performs better in the training process

and poorly in the validation process, it can be concluded that the model is overfit. Consequently, when the test process was carried out, the prediction results were inaccurate.

There are several reasons that could lead the model to overfit, such as imbalance data, or even the data contain noise. Looking at the data proportion table from Table I, it could be clearly seen that the data distribution within the classes were imbalanced as the class Actinic keratosis was only represented by 8% of the total data, and class Dermatofibroma was only represented by 7% of the total data while the other classes were represented by 20% to 30% of the total data. Other than that, when conducting feature extraction using Delaunay triangulation method, it seems that for some images or data, the method could not map the point properly, ultimately giving inappropriate inputs for the model; hence, the terrible accuracy and loss value, and inaccurate predictions or diagnoses.

The last thing to be considered is insufficient features for reliable predictions. The data obtained from the Kaggle website itself has an imbalance data distribution within the classes, with "Pigmented Benign Keratosis" and "Melanoma" as the classes that has the highest number of data. As a result, the model could not learn properly from the classes that has too little data. The differences of the types of skin cancer are challenging to tell even for the human's eye. In this case, the same thing happen for the algorithm. When given to little data to learn from, the algorithm could not classify the type of skin cancer properly. This study paves the way for further in-depth future work to modify the algorithm so that it can perform better and guarantee that the algorithm performs acceptably and predict accurate diagnoses.

REFERENCES

[1] Fox, J. L. (2010). Review: Ultraviolet radiation and skin cancer. *International Journal of Dermatology*, *49*(9). https://onlinelibrary.wiley.com/doi/10.1111/j.1365-4632.2010.04474.x

[2] Apalla, Z. Lallas, A., Sotiriou, E., Lazaridou, E., & Loannides, D. (2017). *Epidemiological trends in skin cancer*. *Dermatology Practical & Conceptual*, *7*(2). https://doi.org/10.5826%2Fdpc.0702a01

[3] MDPI. (2021). Detecting Melanoma Skin Cancer Using Deep Convolutional Neural Networks. Dermatopathology, 8(1), 11. Retrieved from https://www.mdpi.com/2296-3529/8/1/11

[4] ESMO. (2018, May 29). Man Against Machine: Artificial Intelligence is Better than Dermatologists at Diagnosing Skin Cancer. Retrieved from https://www.esmo.org/newsroom/press-releases/artificial-intelligence-skin-cancer-diagnosis

[5] Panja, A., Christy, J. J., & Abdul, Q. M. (2021). An approach to skin cancer detection using Keras and Tensorflow. *Journal of Physics: Conference Series*, 1911(1), 2-3. https://iopscience.iop.org/article/10.1088/1742-6596/1911/1/012032/pdf

[6] Palanichamy, N., Kumar, R. S., Haw, S. C., Ng, K. W., & Anaam, E. (2022). Convolutional neural network models to detect melanoma: a review. *Proceedings of the International Conference on Computer, Information Technology and Intelligent Computing* (CITIC 2022), 470. https://www.atlantis-press.com/proceedings/citic-22/125980678

[7] Zhang, S., Tong, H., Xu, J., & Maciejewski, R. (2019) Graph convolutional networks: a comprehensive review. Computer Social Networks, 6(1). 1. https://www.researchgate.net/publication/337157189_Graph_convolutional_networks_a_comprehensive_review

[8] Yin, S., Peng, Q., Li, H., Zhang, Z., You, X., Liu, H., Fischer, K., Furth, S. L., Tasian, G. E., & Fan, Y. (2019). Multi-instance Deep Learning with Graph Convolutional Neural Networks for Diagnosis of Kidney Diseases Using Ultrasound Imaging. *Bethesda (MD): National Library of Medicine (US)*. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6938161/

[9] Panja, A., Jackson, C. J., & Quadir, A. (2021). An Approach to Skin Cancer Detection Using Keras and Tensorflow. Journal of Physics: Conference Series, 1911(1), 012032. Retrieved from https://doi.org/10.1088/1742-6596/1911/1/012032

[10] Arif, M., Philip, F. M., Ajesh, F., Izdrui, D., Craciun, M. D., & Geman, O. (2022). Automated Detection of Nonmelanoma Skin Cancer Based on Deep Convolutional Neural Network. *Journal of Healthcare Engineering, 2022, Article ID 6952304, 15 pages*. Retrieved from https://doi.org/10.1155/2022/6952304

[11] Katanskiy, A. (2020). *Skin Cancer ISIC*. Retrieved from https://www.kaggle.com/datasets/nodoubttome/skin-cancer9-classesisic

[12] Rotemberg, V., Kurtansky, N., Betz-Stablein, B., Caffery, L., Chousakos, E., Codella, N., Combalia, M., Dusza, S., Guitera, P., Gutman, D., Halpern, A., Helba, B., Kittler, H., Kose, K., Langer, S., Lioprys, K., Malvehy, J., Musthaq, S., Nanda, J., Reiter, O., Shih, G., Stratigos, A., Tschandl, P., Weber, J. & Soyer, P. A patient-centric dataset of images and metadata for identifying melanomas using clinical context. Sci Data 8, 34 (2021). Retrieved from https://doi.org/10.1038/s41597-021-00815-z

[13] Kumar, Y. S., Kumar, N. V., & Guru, D. S. (2015). Delaunay triangulation on skeleton of flowers for classification. *Procedia Computer Science*, *45*, 226-235.

[14] Tzeng, S., & Owens, J. D. (2012). Finding convex hulls using Quickhull on the GPU. *arXiv preprint arXiv:1201.2936*.

[15] Sauras-Altuzarra, L., & Weisstein, E. W. (2007). Adjacency Matrix. In MathWorld--A Wolfram Web Resource. Retrieved from https://mathworld.wolfram.com/AdjacencyMatrix.html

[16] Yao, L., Mao, C., & Luo, Y. (2019). Graph convolutional networks for text classification. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 33, No. 01, pp. 7370-7377)

[17] Arcidiacono, S. (2021, May 5). *What Makes Graph Convolutional Networks Work?* Towards Data Science. Retrieved from https://towardsdatascience.com/what-makes-graph-convolutional-networks-work-53badade0ce9

[18] Oakley, A. (2015, December). *Actinic Keratosis*. DermNet NZ. Retrieved from https://dermnetnz.org/topics/actinic-keratosis

[19] Oakley, A. (2015, December). *Basal Cell Carcinoma*. DermNet NZ. Retrieved from https://dermnetnz.org/topics/basal-cell-carcinoma

[20] Johnson, J. (2021, August 31). *What to Know About Dermatofibromas*. Medical News Today. Retrieved from https://www.medicalnewstoday.com/articles/318870

[21] Heistein, J. B., Acharya, U., & Mukkamalla, S. K. R. (2023). Malignant Melanoma. Treasure Island (FL): StatPearls Publishing.

[22] Thursfield V, Farrugia H, Karahalios E. (2012). Cancer Survival Victoria 2012. Cancer Council Victoria.

[23] MedlinePlus. (2016, August 1). *Epidermal nevus*. Bethesda (MD): National Library of Medicine (US). Retrieved from https://medlineplus.gov/genetics/condition/epidermal-nevus

[24] Singh, S. (2020, June 11). *Data Augmentation to solve imbalanced training data for Image Classification*. Medium. Retrieved from https://medium.com/analytics-vidhya/data-augmentation-to-solve-imbalanced-training-data-for-image-classification-f6d888cbd596

[25] Canuma, P. (2018, October 11). *Image Pre-processing*. Medium. Retrieved from https://prince-canuma.medium.com/image-pre-processing-c1aec0be3edf

[26] Gholamy, A., Kreinovich, V., & Kosheleva, O. (2018). Why 70/30 or 80/20 Relation Between Training and Testing Sets: A Pedagogical Explanation. Departmental Technical Reports (CS), 1209. Retrieved from https://scholarworks.utep.edu/cs_techrep/1209

[27] Suwanda, R., Syahputra, Z., & Zamzami, E. M. (2020, June). Analysis of euclidean distance and manhattan distance in the K-means algorithm for variations number of centroid K. In Journal of Physics: Conference Series (Vol. 1566, No. 1, p. 012058). IOP Publishing.

[28] Doukkali, F. (2018). Batch Normalization in Neural Networks. Retrieved from https://www.kdnuggets.com/2018/06/batch-normalization-neural-networks.htm

[29] Tewari, U. (2021, November 10). Regularization — Understanding L1 and L2 regularization for Deep Learning. Retrieved from https://medium.com/analytics-vidhya/regularization-understanding-l1-and-l2-regularization-for-deep-learning-a7b9e4a409bf

[30] Ide, H., & Kurita, T. (2017, May). Improvement of learning for CNN with ReLU activation by sparse regularization. In 2017 international joint conference on neural networks (IJCNN) (pp. 2684-2691). IEEE.

[31] Farrukh, M. (2019). Modeling on Feature Vectors in Compressed Spaces by the use of Neural network techniques. Retrieved from https://doi.org/10.13140/RG.2.2.29790.59203

[32] Ajagekar, A. (2021). Adam. Retrieved from https://optimization.cbe.cornell.edu/index.php?title=Adam

[33] Ying, X. (2019). An overview of overfitting and its solutions. In *Journal of physics: Conference series* (Vol. 1168, p. 022022). IOP Publishing.