MEKATRONIKA
JOURNAL OF MECHATRONICS AND INTELLIGENT MANUFACTURING
E-ISSN: 2637-0883
VOLUME 6, ISSUE 1, JANUARY 2024, 114 – 119
DOI: https://doi.org/10.15282/mekatronikajintellmanufmechatron.v6i1.10389

RESEARCH ARTICLE

Formulation of A Deep Learning Model for Automated Detection Via Segmentation of Lung Cancer

Liew Yee Zhing[1] and Anwar P.P Abdul Majeed[2]

[1]Faculty of Manufacturing and Mechatronic Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah, 26600 Pekan, Pahang, Malaysia
[2]Department of Computing and Information Systems, Sunway University, Bandar Sunway, 47500 Selangor Darul Ehsan, Malaysia

ABSTRACT - In 2020, cancer accounted for nearly 20 million new cases globally, making it the second leading cause of mortality with over 10 million deaths. In Malaysia alone, there were 48,639 new cases and 29,530 deaths due to cancer in the same year. Among cancers in Malaysia, lung cancer ranks third in frequency but has the highest mortality rate at 15.3 percent, with a mere 11 percent 5-year survival rate. Early detection of lung cancer, often through computed tomography (CT) scanning, is crucial. Pulmonary nodules, small growths on the lung, are indicative of early-stage lung cancer and are typically identified through CT imaging. Distinguishing between benign and malignant nodules is crucial for treatment planning and prognosis. However, manual detection of nodules is labor-intensive and prone to error. Recent advancements in machine learning, particularly transfer learning and fine-tuning techniques, offer promise in automating this process. This study explores the use of transfer learning with ResNet101 and DeepLabV3 for pulmonary nodule segmentation in CT images. Various hyperparameters and neural network architectures were evaluated, with the DeepLabV3-ResNet101-Adagrad Optimizer-Dice Loss pipeline showing the best performance, achieving a Dice Coefficient of 0.7983. These findings hold potential to revolutionize lung cancer screening methods, offering more efficient and accurate detection of pulmonary nodules, ultimately improving patient outcomes.

ARTICLE HISTORY
Received : 28th Feb 2024
Revised : 17th March 2024
Accepted : 30th April 2024
Published : 5th May 2024

KEYWORDS
Lung cancer
Pulmonary modules
Segmentation
Deeplearning
Deeplab V3

1.0 INTRODUCTION

The term "cancer," also known as "malignant tumor" or "malignant neoplasm," refers to a group of disorders characterized by uncontrolled or abnormal cell proliferation. In 2020, the International Agency for Research on Cancer recorded nearly 20 million new cases of cancer around the world [1]. Cancer will be the second biggest cause of mortality worldwide in 2020, with over 10 million deaths. In Malaysia, the recorded number of new cases and deaths due to cancer in 2020 are 48639 and 29530, respectively [2]. Lung cancer is the third most frequent cancer in Malaysia, and it also has the highest mortality rate, at 15.3 percent. Lung cancer has become a major public health issue in Malaysia, with only a 11% 5-year survival rate [3].

Computed tomography (CT) scanning is the most common tool for early-stage lung cancer screening. One of the clinical signs of early lung cancer on CT imaging is pulmonary nodules, which are characterized as a small, opaque, roundish growth on the lung with a size of 7-30mm. There are two types of pulmonary nodules: benign and malignant (cancerous). A doctor or physician can identify the malignancy of lung nodules by the shape and surface of the nodule. Benign nodules have smoother surfaces and a more regular shape, whereas malignant nodules tend to have irregular shapes and rougher surfaces. The characteristic difference between malignant and benign nodules had make the pulmonary nodules segmentation significant as the radiologist can classify the malignancy of the nodules with the size of the nodules. Furthermore, radiologist can adjust the dosage of medication for malignant nodules patient, according to the size of the pulmonary nodules.

In 1966, Lodwick was the first to introduce the notion of computer-aided design (CAD)[4]. Until the 1980s, when relevant theories and technologies were rapidly increasing, CAD technology development came to a halt due to a scarcity of appropriate theories and technologies. The first study on CT-based lung cancer CAD systems was done in 1991. The application of deep learning techniques in pulmonary nodules segmentation has lately become a trend because of the rapid growth of deep learning and its high performance in image recognition.

The general goal of this study is to use computed tomography (CT) images to segment lung nodules. This study will develop different segmentation-based Deep Learning architectures using DeepLabV3 for lung cancer detection from CT images and evaluate the performance of different segmentation-based Deep Learning models in effectively detecting the pulmonary nodules. Several hyperparameter will be tuned to get the best pipeline with best performance in pulmonary

***CORRESPONDING AUTHOR | Anwar P.P Abdul Majeed** | ✉ anwarm@sunway.edu.my

nodule segmentation. The hyperparameter that will be tuned are Epoch, optimizer, and loss function. By using different hyperparameter, the segmentation performance may be varied in CT images segmentation.

1.1 Related Work

Jain et al. in 2021 developed a Salp Shuffled Shepherd Optimization Algorithm-based Generative Adversarial Network (SSSOA-based GAN) in lung nodule segmentation [5]. In the study, the authors used an online database LIDC-IDRI as dataset which consist of 1018 cases CT scans. The CT images is first pre-processed using Gaussian filtering to eliminate the artefacts in the CT images. The pre-processed images are then fed into the Deep Joint segmentation for lung lobe segmentation. The images are then fed into the developed SSSOA-based GAN for lung nodules segmentation. The model is trained by the training set and the performance of the model is evaluated by test set. The generated SSSOA-based GAN model has a maximum Accuracy of 0.9387, a maximum Dice Coefficient of 0.7986, and a maximum Jaccard Similarity of 0.8026, according to the results.

Dutande et al. in 2021 suggested a new method for segmenting, classifying, and detecting lung lesions in CT scan pictures [6]. To do so, the authors proposed SquExUNet segmentation model, 3D-NodNet classification model and 2D-3D cascaded CNN detection model. The proposed SquExUNet model introduced the residual networks and batch normalization into the UNet to captures more information and resolve the problem of vanishing gradient. The dataset in the study is taken from LIDC-IDRI, LNDb Challenge Dataset and Indian Lung CT Image Database (ILCID) which consist of a total number of 1771 CT nodules images. The dataset is then divided in training, validation, and test set. The authors discovered that the suggested model achieved a Dice-Coefficient metrics of 0.80 for nodules segmentation and 90.01 percent Sensitivity for nodules identification by testing the models' performance using a test set.

Chen et al. in 2021 proposed a model called DC-U-Net in lung segmentation of CT images. the model used the U-Net architecture but replace the convolutional layer of the U-Net with dilated convolutional layer [7]. The data comes from the LUNA16 collection, which contains 12,500 CT imaging data sets. The data is separated into two sets, with 60% of the data going to the training set and 40% going to the test set. The train set and test set are then used to train and test the model. The suggested model outperformed the original U-Net with an IOU of 0.9627 and a Dice coefficient of 0.9743, which is much better than the original U-Net.

2.0 METHODS AND MATERIAL

2.1 Data Acquisition and Preprocessing

The Computed Tomography (CT) images for lung cancer patients were retrieved from the public machine learning biomedical image analysis challenge namely "Medical Segmentation Decathlon", created by Antonelli et al. in 2021 [8]. Preoperative thin-section CT images from 96 individuals with non-small cell lung cancer make up the dataset. The lung tumors were the matching target Region of Interest (ROI). The lung tumor label mask is included in the dataset. The difficulty of segmenting tiny areas (tumors) in an image with a broad field-of-view led to the selection of this data set.

The 96 patients with non-small cell lung cancer were divided into three groups: 57 training, 7 validations, and 32 test. Each patient's CT scans may consist of 140 to 150 slices of CT images. All the CT images and labels are stored in NFTI format.

Data normalization is used to convert the values of numeric columns in the dataset to a similar scale without distorting discrepancies in the ranges of values. The first 30 slices CT images of each patient are removed from the dataset as those CT images only show the lower abdomen without lung parenchyma. The CT images and label masks are then resized to 256 x 256 pixel to reduce the computational cost [9].

Data augmentation is then applied to the dataset. The augmentation includes translate images by -15 to +15% on x- and y-axis independently, zoom in or out with scale from 0.85 to 1.15, rotate the images up to 45 degrees and elastic transformation. The oversampling technique is then applied to the training set as the lung tumours are often very small and only appear in few slices among whole CT scans slices. The oversampling can ensure every CT image batch feed into the machine learning model contain at least one CT images with lung tumor.

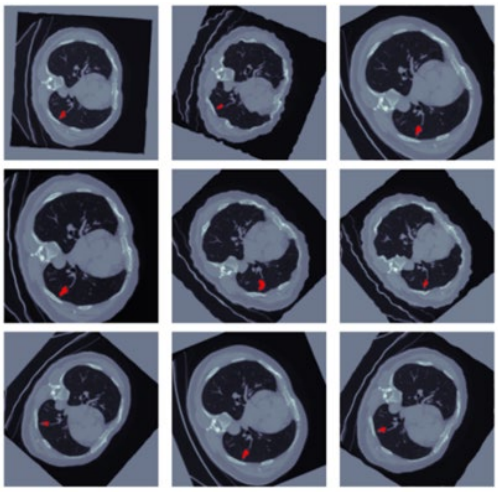


Figure 1. Sample pre-processed images

2.2 Deep Learning Model-DeepLab V3

DeepLabV3 is a semantic segmentation that has numerous improvements over DeepLabV2 [10]. Modules that use atrous convolution in a cascade or parallel fashion to capture multi-scale context by employing multiple atrous rates are designed to address the problem of segmenting objects at many scales. The Atrous Spatial Pyra-mid Pooling (ASPP) module in DeepLabv2 has also been improved with image-level features that encapsulate global context and boost efficiency.

The DeepLabV3 ResNet101 are model built in the PyTorch library. The models are pre-trained on a subset of COCO train2017, on the 20 categories that are present in the Pascal VOC dataset. The DeepLabV3 ResNet101 had achieved mean IOU of 67.4.

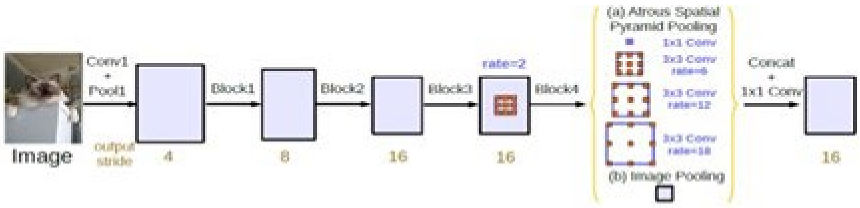


Figure 2. Architecture of DeepLab V3

2.3 Performance Metric

The performance of the models given above is evaluated using performance metrics. The Dice Coefficient and the Intersection of Union (IoU) are two of the performance indicators employed in this study.

The Dice Coefficient is a statistic that is used to compare two samples. Thorvald Srensen (1948) and Lee Raymond Dice (1945) [11], both botanists, developed it separately. A higher Dice Coefficient represents a more robust model. The formula of Dice Coefficient is shown as below:

$$\mathrm{L}(\hat{y}, y) = \mathbf{1} - \frac{\mathbf{2}|\hat{y} \cap y|}{|\hat{y}| + |y|} \tag{1}$$

The Jaccard index, commonly known as the Intersection over Union (IoU) measure, is a method for quantifying the percent overlap between the target mask and forecast output[12]. This metric is like the Dice coefficient, which is frequently used as a training loss function. The IoU metric divides the total number of pixels present across both masks by the number of pixels shared across the target and prediction masks.

$$\mathrm{IoU} = \frac{target \cap \mathrm{prediction}}{target \cup prediction} \tag{2}$$

2.4 Optimizer

Optimizers are algorithms or strategies for minimizing an error function (loss function) or increasing production efficiency. Optimizers are mathematical functions that are based on learnable parameters in a model, such as Weights and Bias. Three distinct optimizers are utilized in this study to discover the optimal pipeline for segmenting pulmonary nodules. Adam Optimizer, SGD Optimizer, and Adagrad Optimizer are the optimizers used.

Adaptive Moment Estimation is a technique for optimising gradient descent algorithms [12] When working with huge problems with a lot of data or parameters, the method is quite efficient. It is efficient and takes minimal memory. It appears to be a hybrid of the 'gradient descent with momentum' and the 'RMSP' algorithms. Adam Optimizer takes the best features of the previous two methods and improves on them to produce a more optimized gradient descent.

The iterative method of stochastic gradient descent (commonly abbreviated SGD) for maximizing an objective function with sufficient smoothness criteria (e.g. differentiable or subdifferentiable) [13]. Because it replaces the actual gradient (derived from the complete data set) with an estimate, it can be considered a stochastic approximation of gradient descent optimization (calculated from a randomly selected subset of the data). This minimizes the extremely high computational cost, especially in high-dimensional optimization problems, allowing for faster iterations in exchange for a reduced convergence rate.

Because there is no idea of momentum in the Adagrad optimizer, it is much simpler than SGD with momentum [14]. Adagrad's concept is to apply various learning rates for each parameter based on iteration. The necessity for variable learning rates arises from the fact that sparse features parameters require a greater learning rate than dense features parameters due to the lower frequency of occurrence of sparse features.

2.5 Loss Function

The loss function is a function that calculates the difference between the algorithm's current output and the expected output. It's a tool for assessing how well your algorithm models data. If the forecasts are incorrect, the loss function will return a greater value. If they're decent, it'll give you a lower number. Your loss function will inform you if you're making progress as you tweak parts of your algorithm to try to enhance your model. Three different loss functions are employed in this study to find the optimum pipeline. Dice Loss, Binary Cross Entropy Loss, and Focal Loss are the three.

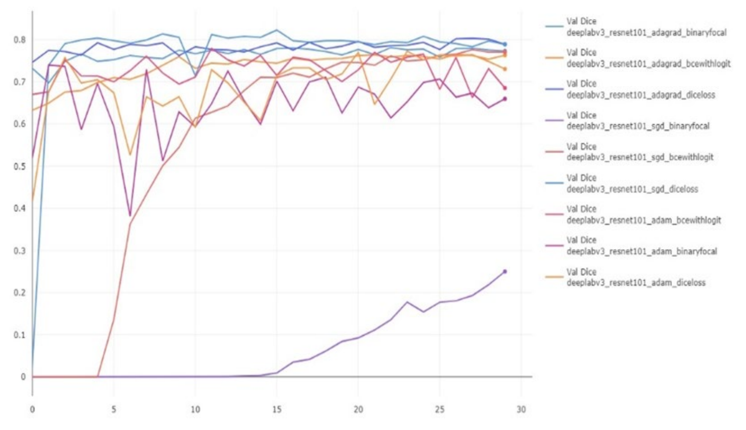
The term dice loss comes from the Srensen–Dice coefficient, a statistic created in the 1940s to determine how comparable two samples are. In 2016, it was introduced to the computer vision community for the purpose of 3D medical image segmentation. Dice Loss is just 1 - Dice Coefficient, as previously stated. As a result, Dice loss analyses both local and global loss information, which is crucial for high accuracy.

When a classification problem only has two categories, the loss function Binary Cross Entropy is utilized. Its name implies that it represents two quantities, which is why it is built in such a way that it meets the problem of classification of two quantities. Because we have a binary segmentation problem, Binary Cross Entropy is also appropriate in this study. Because this function is utilized with discrete data, the Probability Mass Function (PMF) (return probability) is employed in this situation instead of the Probability Density Function (PDF) (return density) in the case of continuous values when Mean Squared Error was used.

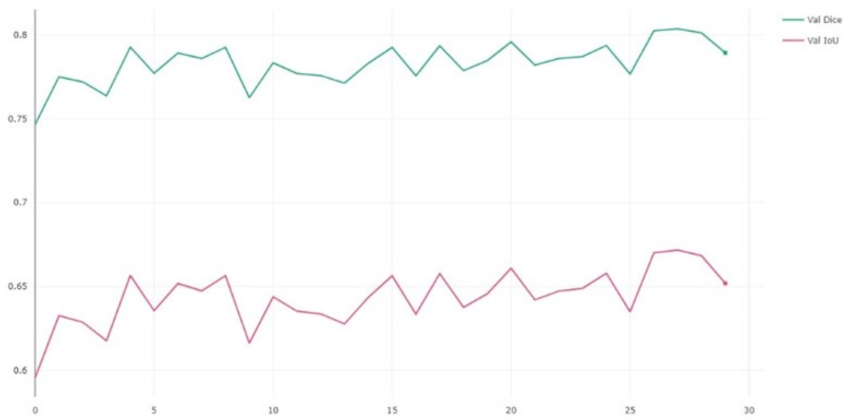
During object detection training, a Focal Loss function tackles class imbalance. To focus learning on hard misclassified examples, focal loss adds a modifying term to the cross-entropy loss. It's a dynamically scaled cross entropy loss, in which the scaling factor decreases as confidence in the proper class grows. This scaling factor, intuitively, can automatically down-weight the contribution of easy samples during training and quickly focus the model on difficult situations.

3.0 RESULTS AND DISCUSSION

From the result DeepLabV3+ResNet101+Adagrad+DiceLoss pipeline are the best performing pipeline with a Dice Coefficient of 0.7893. The Dice Coefficient graph over Epoch of both pipelines are shown in the Figure 3 (a) and (b) below. From the figure, it is clearly seen that the pipeline has a high performance on segmentation of lung cancer at the same time had low fluctuation over the epoch during the training. The training time of both pipeline is short too. With a minimal training epoch of 10, the model able to converge to get good result.



(a)



(b)

Figure 3. **(a)** Dice coefficient of all pipeline **(b)** deeplab V3+ResNet101+Adagrad+DiceLoss pipeline

Table 1. Performance of all pipeline with highlighted best performance pipeline

Pipeline (Model+Encoder+Optimizer+Loss function)	Dice score	IoU
DeepLabV3+ResNet101+Adam+DiceLoss	0.730	0.575
DeepLabV3+ResNet101+Adam+BinaryCrossEntropyLoss	0.685	0.521
DeepLabV3+ResNet101+Adam+FocalLoss	0.659	0.492
DeepLabV3+ResNet101+Adagrad+DiceLoss	0.789	0.651
DeepLabV3+ResNet101+Adagrad+BinaryCrossEntropyLoss	0.763	0.617
DeepLabV3+ResNet101+Adagrad+FocalLoss	0.773	0.630
DeepLabV3+ResNet101+SGD+DiceLoss	0.788	0.651
DeepLabV3+ResNet101+SGD+BinaryCrossEntropyLoss	0.770	0.627
DeepLabV3+ResNet101+SGD+FocalLoss	0.250	0.142

DeepLabV3+ResNet101+SGD+FocalLoss is the worst performing pipeline in segmentation of lung tumor with dice coefficient and IoU of 0.25 and 0.1428 respectively. From the result, the combination of SGD optimizer and Focal Loss loss function leads to a bad result. From Figure 4 the SGD+FocalLoss pipeline required more time or epoch to start converging. The model start converging at 15th epoch.

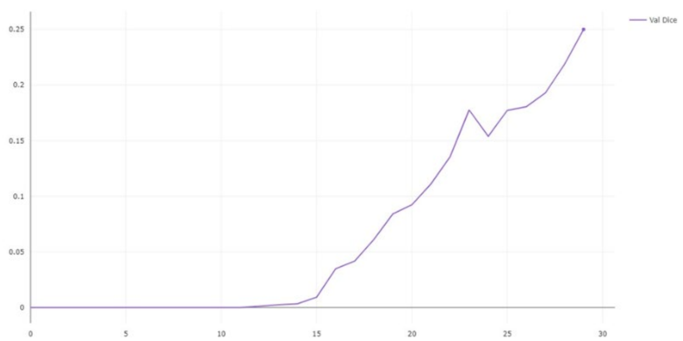


Figure 4. Dice coefficient of pipeline Deeplab V3+ResNet101+SGD+FocalLoss

4.0 CONCLUSION

In conclusion, this research is successfully conducted and can answer the research questions early on this paper. Throughout the study, few hyperparameter had been tuned for the deep learning model DeepLabV3 in segmentation of lung cancer. The hyperparameter include encoder, epoch, optimizer, and loss function. Two performance metrics had been used in the study which there are Dice Coefficient and IoU. From the study, the DeepLabV3+ResNet101+Adagrad+DiceLoss that used ResNet101 as encoder yield the best result in segmentation of lung cancer. The model achieved a Dice Coefficient of 0.7983 and IoU of 0.6519.

The scope of this study was limited in terms of evaluating other hyperparameters and factors. The following recommendations are suggested in extending and explore the present study. The present study utilizes dataset from United Stated, but the CT images of Asian and American may varies. Therefore, future study should consider using other dataset from Malaysia to show the effectiveness of the model in segmentation of Asian lung cancer.

The author understands the scarcity of dataset for lung cancer CT images. Therefore, hyperparameter tuning is best option in increasing the segmentation performance with a limited number of dataset available. Future studies should investigate hyperparameters other than mentioned in the present study to further decrease the mis-segmentation in the present literature.

5.0 ACKNOWLEDGMENT

The authors would like to thank Universiti Malaysia Pahang Al-Sultan Abdullah for the funding of this study.

6.0 REFERENCES

[1] "Cancer Today." Accessed: Apr. 19, 2024. [Online]. Available: https://gco.iarc.fr/today/en

[2] P. K. Nureylia Amir Nabihah Ali Sukumaran Raman Balqis Bahtiar Renuga Raman, *NATIONAL CANCER REGISTRY DEPARTMENT National Cancer Institute, Ministry Of Health MALAYSIAN STUDY ON CANCER SURVIVAL (MySCan) National Cancer Institute Ministry of Health Putrajaya MALAYSIA Health Education Division Ministry of Health Putrajaya MALAYSIA i Malaysian Study On Cancer Survival (MySCan), 2018 MINISTRY OF HEALTH*. 2014. [Online]. Available: http://nci.moh.gov.my/index.php/ms/main-menu-2/laporan

[3] "TEN MOST COMMON CANCERS IN MALAYSIA," 2007.

[4] Y. Gu *et al.*, "A survey of computer-aided diagnosis of lung nodules from CT scans using deep learning," *Comput Biol Med*, vol. 137, Oct. 2021, doi: 10.1016/J.COMPBIOMED.2021.104806.

[5] S. Jain, S. Indora, and D. K. Atal, "Lung nodule segmentation using Salp Shuffled Shepherd Optimization Algorithm-based Generative Adversarial Network," *Comput Biol Med*, vol. 137, Oct. 2021, doi: 10.1016/J.COMPBIOMED.2021.104811.

[6] P. Dutande, U. Baid, and S. Talbar, "LNCDS: A 2D-3D cascaded CNN approach for lung nodule classification, detection and segmentation," *Biomed Signal Process Control*, vol. 67, May 2021, doi: 10.1016/J.BSPC.2021.102527.

[7] K. bing Chen, Y. Xuan, A. jun Lin, and S. hua Guo, "Lung computed tomography image segmentation based on U-Net network fused with dilated convolution," *Comput Methods Programs Biomed*, vol. 207, Aug. 2021, doi: 10.1016/J.CMPB.2021.106170.

[8] M. Antonelli *et al.*, "The Medical Segmentation Decathlon," *Nat Commun*, vol. 13, no. 1, Jun. 2021, doi: 10.1038/s41467-022-30695-9.

[9] R. C. Hardie, A. T. Trout, J. R. Dillman, B. N. Narayanan, and A. A. Tanimoto, "Performance Analysis in Children of Traditional and Deep Learning CT Lung Nodule Computer-Aided Detection Systems Trained on Adults," *American Journal of Roentgenology*, vol. 222, no. 2, Feb. 2024, doi: 10.2214/AJR.23.30345.

[10] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking Atrous Convolution for Semantic Image Segmentation," Jun. 2017, Accessed: Apr. 19, 2024. [Online]. Available: https://arxiv.org/abs/1706.05587v3

[11] D. P. Kingma and J. L. Ba, "Adam: A Method for Stochastic Optimization," *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, Dec. 2014, Accessed: Apr. 19, 2024. [Online]. Available: https://arxiv.org/abs/1412.6980v9

[12] S. Ruder, "An overview of gradient descent optimization algorithms," Sep. 2016, Accessed: Apr. 19, 2024. [Online]. Available: https://arxiv.org/abs/1609.04747v2

[13] J. C.Duchi, E. Hazan, and Y. Singer, "(PDF) Adaptive Subgradient Methods for Online Learning and Stochastic Optimization," Journal of Machine Learning. Accessed: Apr. 19, 2024. [Online]. Available: https://www.researchgate.net/publication/220320677_Adaptive_Subgradient_Methods_for_Online_Learning_and_Stochastic_Optimization

[14] L. R. Dice, "Measures of the Amount of Ecologic Association Between Species," *Ecology*, vol. 26, no. 3, pp. 297–302, Jul. 1945, doi: 10.2307/1932409.