NATURENGS
MTU Journal of Engineering and Natural Sciences
https://dergipark.org.tr/tr/pub/naturengs
DOI: 10.46572/naturengs.1205919

Research Article

Detection of Tonsillopharyngitis with Grad-Cam and Optimization-Based Model Using Oropharyngeal Images

Muhammed YILDIRIM[1*], Orkun EROGLU[2]

[1] Department of Computer Engineering, Faculty of Engineering and Natural Sciences, Malatya Turgut Ozal University, Malatya, Turkey.
[2] Department of Otorhinolaryngology, School of Medicine, Firat University, Elazig, Turkey.

(**Received:** 17.11.2022; **Accepted:** 09.12.2022)

ABSTRACT: Tonsillopharyngitis is a sudden onset contagious infection of the pharynx and tonsils. Patients experience a rapid general condition and loss of workforce. In addition to affecting patients, it spreads and affects other individuals. In addition, it causes severe complications and increases hospital costs. Therefore, early and accurate diagnosis is essential. In this study, a hybrid model is developed to diagnose tonsillopharyngitis. First, the heat maps of the images in the original data set by applying the Gradient-weighted Class Activation Mapping (Grad-Cam) method. In the proposed model, feature maps are obtained from the original and heatmap datasets using the Darknet53 architecture as the base. It aims to increase the proposed model's performance by bringing together different features of the same image. After the feature map obtained after the feature fusion step is optimized with the Relief method, classification is carried out using an SVM shallow classifier.

Keywords: Artificial Intelligence, CNN, Grad-Cam, Relief, Tonsillopharyngitis.

1. INTRODUCTION

Tonsillopharyngitis is a contagious infection and inflammation of the pharynx and tonsils. It is one of the most common infections in children and adults [1]. Bacteria and viruses can cause infection. It is one of the common reasons for hospital admissions, especially in childhood. Transmission is more common in communal living areas. Tonsillopharyngitis caused by microorganisms called group A beta-hemolytic streptococcus is most common in school-age children. The disease often increases in the winter season. The most common complaints of the patients are sudden onset of sore throat, fever, and inability to swallow and feed. Physical examination, throat culture, serological tests, and rapid antigen tests are used for diagnosis. Physical examination may reveal swelling of the tonsils, painful lymphadenopathy in the anterior part of the neck, and rash on the soft palate. Early diagnosis of the disease prevents the development of complications in patients as early treatment will be initiated. In addition, early diagnosis prevents the transmission of the disease to other people [2-4].

Tonsillopharyngitis is a common life-threatening infection when complications develop. It is a significant health problem because it causes loss in the workforce and can easily be transmitted in family and communal living spaces. Detection of this problem at an early stage is of great

[*]Corresponding Author: muhammed.yildirim@ozal.edu.tr
ORCID number of authors: [1] 0000-0003-1866-4721, [2] 0000-0001-9392-5755

importance for treating the patient [5]. In recent years, the use of artificial intelligence methods in the medical field has been the focus of the attention of researchers [6-8]. This study aims to recognize tonsillopharyngitis images by developing an artificial intelligence-based hybrid model.

The novelties and contributions of the study are as follows.

- In this study, heat maps of oropharynx images in the data set were obtained with Grad-Cam technology.
- Feature maps were extracted from the original and heatmap datasets using the Darknet53 architecture. In this step, different features of the same image are obtained.
- It aims to increase the model's performance by combining the features obtained from both the original data set and the data set created by obtaining heat maps.
- The Relief method is used to reduce the size of the feature map obtained after the feature merging step. In this step, it is aimed that the model will run faster and produce more successful results since the unnecessary features are eliminated and the size reduction process is performed.
- High performance is achieved by classifying the optimized feature map in the SVM shallow classifier.
- While the highest accuracy value obtained in eight different pre-trained models is 84.71%, the accuracy value obtained in the proposed model is 89.7%.

Not many studies were found when the literature on the subject was searched. Yoo et al. study created a data set consisting of 131 throats and 208 normal throat images. Since the images in the created data set were few, the researchers made data multiplexing. GAN structures were used for data multiplexing. In the study, researchers used ResNet50, InceptionV3 and MobileNetV2 architectures. The accuracy value obtained by using 4-fold cross-validation in the ResNet50 model was 95.3%. The researchers stated that the proposed system could be used in smartphone-based applications [9].

The article's organization is as follows: The material and method part is examined in the second part. In the third part, the results of the application are presented in detail. The fourth part of the study, discussion, and conclusions in the fifth part are included.

2. MATERIAL AND METHODS

This section examines the data set used in the study, pre-trained models, classifiers, Grad-Cam technology, and the proposed model. The proposed model aims to obtain high success rates in classifying tonsillopharyngitis images.

2.1. Data set

Tonsillopharyngitis is a common, acute-onset infection of the pharynx and tonsils. Patients have sore throats and difficulty swallowing. Physical examination is essential in diagnosis. Physical study reveals redness in the oropharynx, swelling of the tonsils, and exudate. The infection may spread to the surrounding soft tissues in the neck and cause serious complications such as peritonsillar abscesses and mild neck tissue infections. The data set used in the study consists of oropharyngeal images with tonsillopharyngitis and normal oropharyngeal images. While creating this data set we used, the researchers applied the knowledge of two experts. In addition, ambiguous images were eliminated by experts. While collecting the data set used in the study, images up to April 2020 were collected [9, 10]. The related data set contains 131 images of the oropharynx with tonsillopharyngitis and 208 images of the normal oropharynx. Image examples from the original dataset are presented in Figure 1.

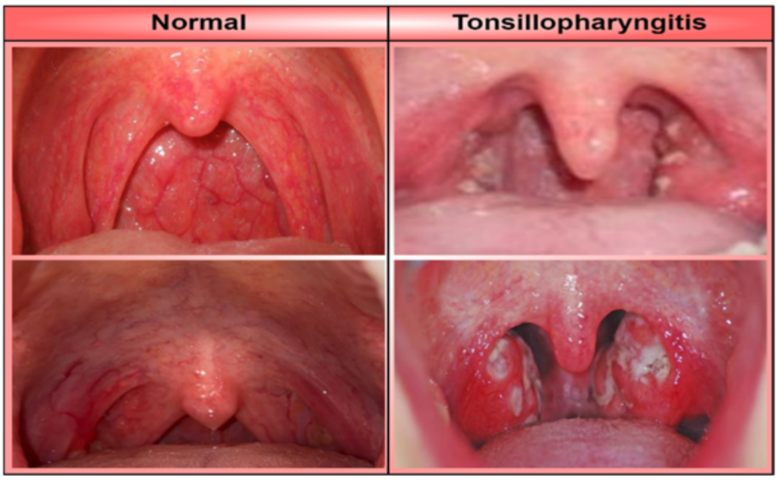

Figure 1. Examples from oropharynx images

2.2. Grad-Cam technology

Grad-Cam technology is a technique used to obtain heat maps of images. Grad-Cam uses gradients of any target concept. Grad-Cam can be applied to various CNN architectures [11]. In this study, heat maps of existing images were obtained using Grad-Cam technology. Then, the feature maps of these images were extracted and combined with the feature maps of the original images. This way, the model's performance is increased by using different features of the same image. The image samples obtained by applying Grad-Cam to the original images are given in Figure 2.

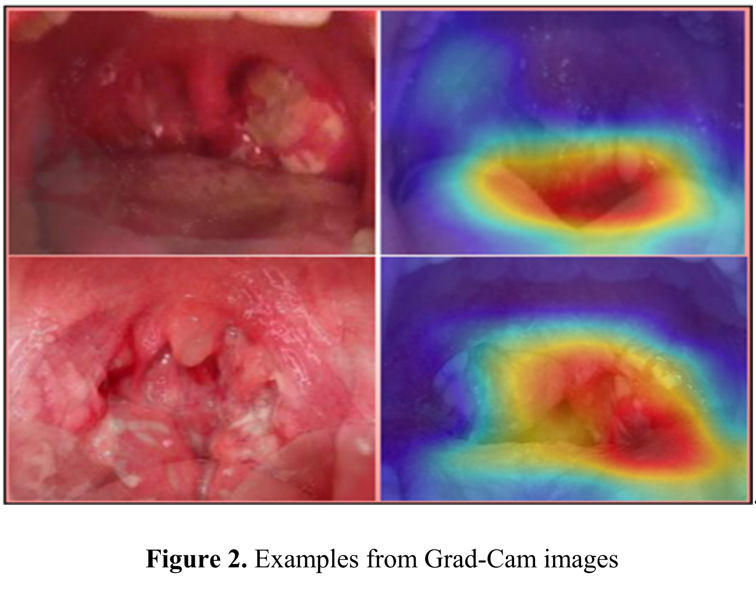

Figure 2. Examples from Grad-Cam images

2.3. Pre-trained architectures, Relief, and classifiers

Artificial intelligence-based methods have been used frequently in recent years, especially in the biomedical field. Especially after AlexNet architecture won the ImageNet competition in 2012, deep learning, which had a stagnant period, started to become popular again. Eight different pre-trained architectures were used in this study to diagnose tonsillopharyngitis using oropharyngeal images. The results obtained in these architectures used in the literature were compared with the proposed hybrid model. In this study, AlexNet [12], DarkNet53 [13], DenseNet201 [14], EfficientNetb0 [15], GoogleNet [16], MobileNetV2 [17], ResNet101 [18], and ShuffleNet [19] architectures were used. In these CNN architectures used in the study, feature extraction is done automatically. In these architectures, there is no need for expert knowledge in the feature extraction stage.

After combining the feature maps obtained with different techniques in the proposed hybrid model, the Relief method optimized the obtained feature map [20]. After eliminating unnecessary features, the optimized feature map obtained was classified into different machine learning classifiers. The classifiers used in the study are Support Vector Machine (SVM)[21], k-nearest neighbor (KNN)[22], Ensemble Subspace (ES)[23], Decision Tree(DT)[24], Discriminant Analysis (DA)[25], Naïve Bayes (NB)[26] and Logistic Regression(LR)[27].

2.4. Developed model

In the model developed for the classification of oropharyngeal images, first of all, heat maps of the pictures of the original data set were obtained with Grad-Cam technology. Then, feature maps were extracted from both the original dataset and the dataset created using Grad-Cam technology using DarkNet53. After these extracted feature maps were combined, the feature map was optimized using Relief. After the Relief eliminated the unnecessary features, these features were classified into different classifiers. In the proposed model, the highest success rates were obtained in SVM.

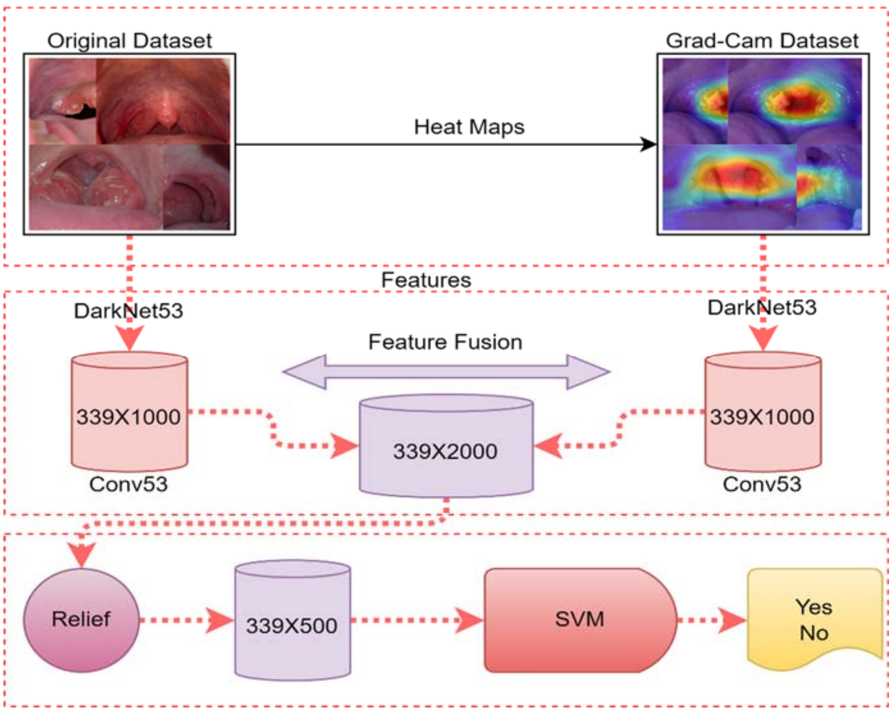

Figure 3. Proposed model for the classification of oropharyngeal images

As seen in Figure 3, feature extraction was done with DarkNet53 architecture. Features are taken from the Conv53 layer of the DarkNet53 architecture. 1000 features were taken for each image from the original data set and the data set created with Grad-Cam technology. The dimensions of the feature maps obtained from the original dataset and with Grad-Cam technology are 339x1000. After the merge, the size of the new feature map is 339x2000. For the model developed for classifying oropharyngeal images to work faster and more efficiently, the features in the size of 339x2000 were optimized with the Relief method. The size of the feature map optimized by the Relief method was 339x500. The optimized feature map in the last step is classified in the SVM classifier. The results were obtained by using the pre-trained models in the literature to compare the performances of the developed model.

3. RESULTS

In this study, which was carried out to diagnose tonsillopharyngitis using oropharyngeal images, the results of the application were obtained in a Matlab environment. To compare the performance of the proposed hybrid model, results were obtained with 8 pre-trained models accepted in the literature using the original data set. While training the pre-trained models, 75% of the images in the original dataset were used. The remaining 25% of the dataset is reserved for testing the models. In addition, the epoch value of 5, the InitialLearnRate value of 1e-4, and the MiniBatchSize value of 16 were selected while training the pre-trained models.

The hybrid model developed for classifying oropharyngeal images was also used as the basis of the DarkNet53 architecture, and feature extraction was made from both the original data set and the data set obtained using Grad-Cam technology. Then, the obtained features were combined and optimized and classified in different classifiers. In this way, each of the models used in the study was tested with a further 20% data set each time. Default training parameters and 5-fold cross-validation were used in shallow classifiers.

Different evaluation metrics were used to measure the performance of the models used in the study to diagnose tonsillopharyngitis disease using oropharyngeal images. Accuracy, Sensitivity, Specificity, Negative Predictive Value(NPV), False Positive Rate(FPR), False Negative Rate(FNR), False Discovery Rate(FDR), F1 score, and Matthews Correlation Coefficient(MCC) are the leading performance measurement metrics used in the study [28-30].

3.1. Results of pre-trained models

This study obtained results primarily in pre-trained models for diagnosing tonsillopharyngitis using oropharyngeal images. The results obtained in these models were compared with the developed model. The confusion matrix of the pre-trained models is in Figure 4.

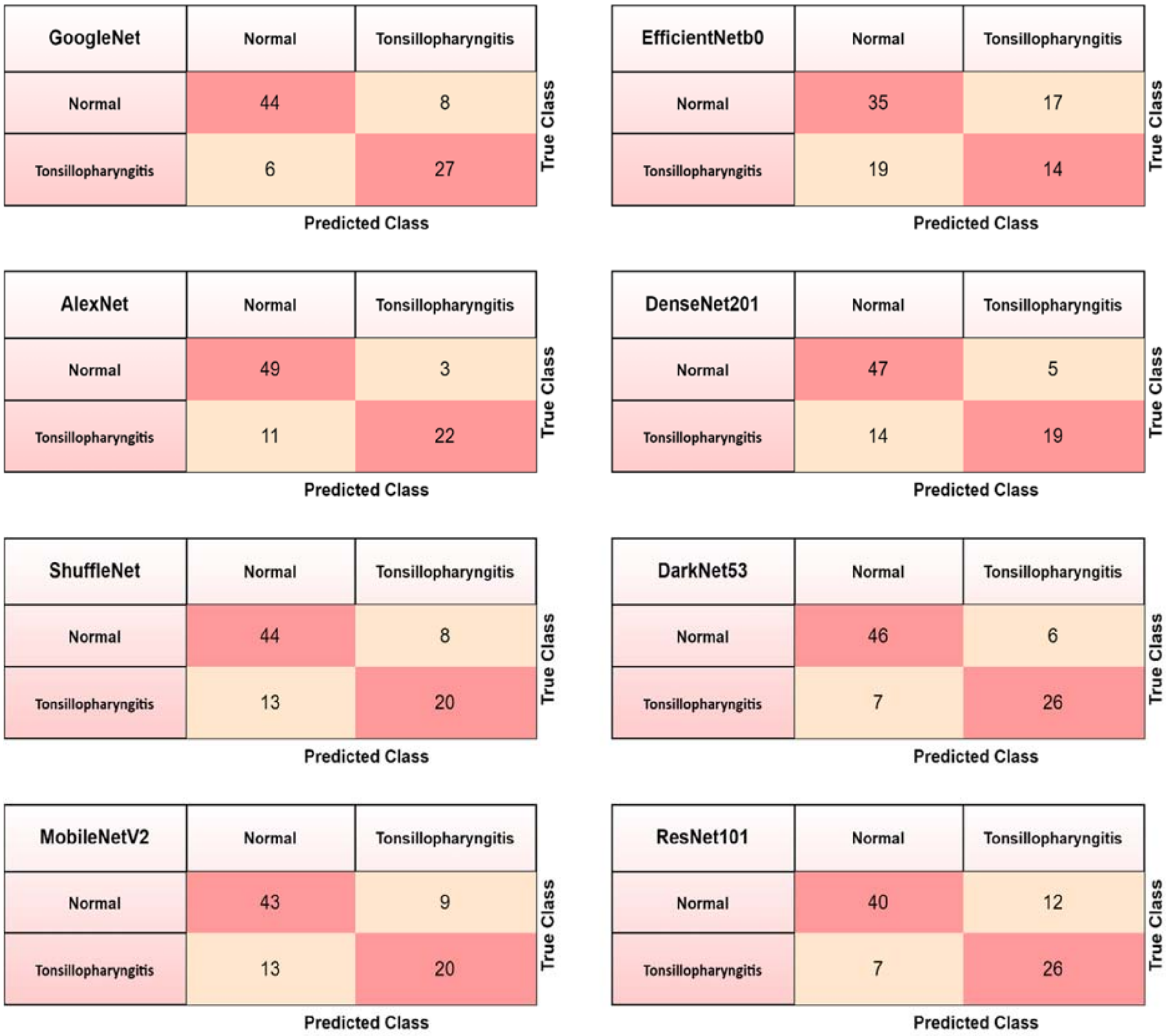

Figure 4. Confusion matrix of the pre-trained models

The EfficientNetb0 architecture predicted 49 correctly and 36 incorrectly of the 85 images used for testing. The success of this model in classifying oropharyngeal images was 57.64%. The second model used to diagnose tonsillopharyngitis using oropharyngeal images is MobileNetV2. The MobileNetV2 architecture predicted 63 correctly and 22 incorrectly of the 85 images used for testing. The success of this model in classifying oropharyngeal images was 74.11%. The third model used to diagnose tonsillopharyngitis using oropharyngeal images is ShuffleNet. This architecture predicted 64 correctly and 21 incorrectly of the 85 images used for testing. The success of this model in classifying oropharyngeal images is 75.29%. The fourth model used to diagnose tonsillopharyngitis using oropharyngeal images is the ResNet101. The ResNet101 predicted 66 correctly and 19 incorrectly of the 85 images used for testing. The success of this model in classifying oropharyngeal images is 77.64%. The fifth model used to diagnose tonsillopharyngitis using oropharyngeal images is DenseNet201. The DenseNet201 predicted 66 correctly and 19 incorrectly of the 85 images used for testing. The success of this model in classifying oropharyngeal images is 77.64%. The sixth model used to diagnose tonsillopharyngitis using oropharyngeal images is the AlexNet. The AlexNet predicted 71 correctly and 14 incorrectly of the 85 images used for testing. The success of this model in classifying oropharyngeal images is 83.52%. The seventh model used to diagnose tonsillopharyngitis using oropharyngeal images is GoogleNet. The GoogleNet predicted 71 correctly and 14 incorrectly of the 85 images used for testing. The success of this model in classifying oropharyngeal images is 83.52%. The last model used to study is the DarkNet53. The DarkNet53 predicted 72 correctly and 13 incorrectly of the 85 images used for testing. The success of this model in classifying oropharyngeal images is 84.71%. The results of the pre-trained models used in classifying oropharyngeal images are in Table 1.

Table 1. Accuracy of pre-trained models

Model	Accuracy
EfficientNetb0	57.64%
MobileNetV2	74.11%
ShuffleNet	75.29%
ResNet101	77.64%
DenseNet201	77.64%
AlexNet	83.52%
GoogleNet	83.52%
DarkNet53	84.71%

When the results obtained in 8 different pre-trained models on oropharyngeal images for the diagnosis of Tonsilopharyngitis disease are examined in Table 1, it is seen that the highest accuracy rate is obtained in the DarkNet53 architecture. DarkNet53 is the most accurate model among the pre-trained models for classifying oropharyngeal images.

3.2. The results of the CNN architectures used as the base in the proposed model

In this section, DarkNet53 architectures are used as the basis and the results obtained in 7 different classifiers are examined. First, a new data set was created by applying the Grad-Cam method to the original data set, for which heat maps were obtained. In the next step, DarkNet53 architectures were used as the base, and feature maps were extracted from the two existing datasets. After these extracted feature maps were combined, unnecessary features were eliminated by the Relief method. After the Relief method, features decreased from 2000 to 500. In the last step, the feature map optimized with the Relief method was classified into 7 different shallow classifiers. The obtained accuracy values are presented in Table 2.

Table 2. Accuracy Rate of Proposed Models at Different Classifiers (%)

	DT	DA	LR	SVM	NB	KNN	SE
Proposed Model	72.0	72.9	55.2	**89.7**	79.1	87	86.1

In the last step of the proposed model, accuracy values were obtained using different classifiers. Among the 7 different classifiers used in the study, the highest accuracy value was obtained in the SVM classifier at 89.7%. In contrast, the lowest accuracy value was obtained in the LR classifier at 55.2%. The confusion matrix obtained proposed model for classifying oropharyngeal images is presented in Figure 5.

Proposed Model	Normal	Tonsillopharyngitis
Normal	195	13
Tonsillopharyngitis	22	109

Predicted Class

True Class

Figure 5. Confusion matrix of the proposed model for classification of oropharyngeal images

When Figure 5 is examined, it is seen that the proposed model for the classification of oropharyngeal images predicts 304 of 339 oropharyngeal images correctly and 35 of them incorrectly. It is seen that the proposed model predicts 195 of 208 Normal oropharyngeal images correctly and 13 of them as Tonsilopharyngitis. Similarly, out of 131 oropharyngeal images belonging to the Tonsilopharyngitis class, the proposed model predicted 109 correctly and 22 incorrectly. The success rate of the proposed model in classifying oropharyngeal images is 89.7%. The success metrics of the proposed model in classifying oropharyngeal images are in Table 3.

Table 3. Performance metrics of the proposed model for the classification of oropharyngeal images

Accuracy	Sensitivity	Specificity
89.7%	89.9%	89.3%
NPV	**FPR**	**FNR**
83.2%	0.1066	0.1014
FDR	**F1 score**	**MCC**
0.0625	91.8%	78.1%

When the performance evaluation metrics given in Table 3 are examined, it is seen that the proposed model for the classification of oropharyngeal images achieves high performance in the category of oropharyngeal images.

4. DISCUSSION

Tonsillopharyngitis is a common life-threatening infection when complications develop. It is an infection of the tonsil and pharynx tissues caused by bacteria or viruses. Chronic infection of the tonsils and obstructive hypertrophic tonsils are among the most common diseases in the pediatric age group [31]. It causes many complications, such as mouth breathing, nasal speech, maxillofacial developmental disorders, and psychosocial, neurocognitive, and developmental disorders in the following years [31]. In addition, serious complications such as peritonsillar abscess, deep neck soft tissue infection, and arthritis may be seen in patients whose treatment is delayed [32].

Tonsillopharyngitis causes a sudden onset of addiction, resulting in loss of workforce. In addition, it is a significant health problem since it can be easily transmitted within the family and in public places. For this reason, early and accurate diagnosis is essential because the disease can spread rapidly and cause serious complications with high mortality if the treatment is delayed. In this study, tonsillopharyngitis was diagnosed using artificial intelligence methods in the data set consisting of tonsillopharyngitis and normal oropharyngeal images. High success rates have been achieved in the hybrid model developed for this purpose. The developed hybrid model was also compared with the pre-trained models in the literature. When the results obtained in the developed model are compared with the results obtained in the pre-trained models, it is seen that there is a significant increase.

The accuracy values of the pre-trained models and the proposed model used in the study for the classification of oropharyngeal images are compared in Figure 6.

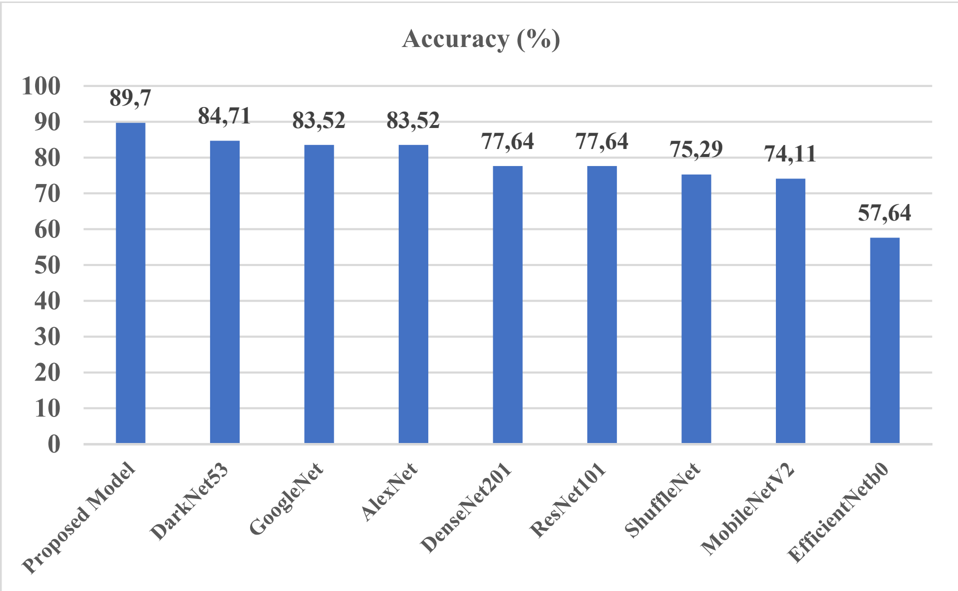

Figure 6. Compare of model's accuracy rate

When Figure 6 is examined, the proposed model's highest accuracy value is 89.7%. The closest value to this value was obtained in DarkNet53 with 84.71%.

This study used the CNN-based DarkNet53 model as the basis for feature extraction. Heat maps of the images in the data set were obtained using Grad-Cam technology. Features were extracted from the data sets created using Grad-Cam and interpolation method and these features were combined. The Relief dimension reduction method performed feature reduction after the feature fusion process. The number of features has been reduced from 2000 to 500. In this way, it is aimed to run the model faster. The optimized features in the last step of the model are classified in the SVM classifier.

The study's biggest limitation is that the data is obtained from a single center. Testing the model with data obtained from different centers would be more appropriate. It is among our aims to get data from various centers and work with more experts.

5. CONCLUSIONS

Tonsillopharyngitis is a disease with high morbidity and mortality. Detection of this disease by computer-aided systems is of great importance. Computer-assisted artificial intelligence-based systems have been used frequently in recent years, especially in the biomedical field, and successful results have been obtained. In this study, tonsillopharyngitis was diagnosed with an accuracy rate of 89.7% in the data set consisting of oropharyngeal images. This value shows that the proposed computer-assisted artificial intelligence-based model can be used to diagnose Tonsillopharyngitis.

Acknowledgments
The authors thank the owners of the database for sharing their data.

Declaration of Competing Interest
The authors declare that there is no conflict of interest in the study.

Author Contribution

The authors contributed equally to the article.

REFERENCES

[1] Bochner, R.E., Gangar, M., Belamarich, P.F., A clinical approach to tonsillitis, tonsillar hypertrophy, and peritonsillar and retropharyngeal abscesses. Pediatrics in Review, 2017. 38(2): p. 81-92.

[2] Vicedomini, D., et al., Diagnosis and management of acute pharyngotonsillitis in the primary care pediatrician's office. Minerva Pediatrica, 2014. 66(1): p. 69-76.

[3] Demir, N., Bayar Muluk, N., Chua, D., Acute Tonsillopharyngitis in Children, in Pediatric ENT Infections. 2022, Springer. p. 515-523.

[4] Osiejewska, A., et al., Acute tonsillopharyngitis-a review. Journal of Education, Health and Sport, 2022. 12(7): p. 873-882.

[5] Amiraraghi, N., et al., Intramural oesophageal abscess: an unusual complication of tonsillitis. BMJ Case Reports CP, 2019. 12(2): p. bcr-2018-226010.

[6] Bingol, H., NCA-based hybrid convolutional neural network model for classification of cervical cancer on gauss-enhanced pap-smear images. International Journal of Imaging Systems and Technology, 2022.

[7] Toğaçar, M., B. Ergen, Tümen, V., Use of dominant activations obtained by processing OCT images with the CNNs and slime mold method in retinal disease detection. Biocybernetics and Biomedical Engineering, 2022.

[8] Kiziloluk, S. and Sert, E., COVID-CCD-Net: COVID-19 and colon cancer diagnosis system with optimized CNN hyperparameters using gradient-based optimizer. Medical and Biological Engineering & Computing, 2022. 60(6): p. 1595-1612.

[9] Yoo, T.K., et al., Toward automated severe pharyngitis detection with smartphone camera using deep learning networks. Computers in biology and medicine, 2020. 125: p. 103980.

[10] https://data.mendeley.com/datasets/8ynyhnj2kz/1.

[11] Selvaraju, R.R., et al., Grad-cam: Visual explanations from deep networks via gradient-based localization. in Proceedings of the IEEE international conference on computer vision. 2017.

[12] Krizhevsky, A., Sutskever, I., Hinton, G.E., Imagenet classification with deep convolutional neural networks. Communications of the ACM, 2017. 60(6): p. 84-90.

[13] Redmon, J., Farhadi, A., Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767, 2018.

[14] Huang, G., et al. Densely connected convolutional networks. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

[15] Tan, M., Le, Q., Efficientnet: Rethinking model scaling for convolutional neural networks. in International conference on machine learning. 2019. PMLR.

[16] Szegedy, C., et al. Going deeper with convolutions. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.

[17] Sandler, M., et al. Mobilenetv2: Inverted residuals and linear bottlenecks. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

[18] He, K., et al. Deep residual learning for image recognition. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[19] Zhang, X., et al. Shufflenet: An extremely efficient convolutional neural network for mobile devices. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

[20] Yu, L., Liu, H., Feature selection for high-dimensional data: A fast correlation-based filter solution. in Proceedings of the 20th international conference on machine learning (ICML-03). 2003.

[21] Joachims, T., 11 making large-scale support vector machine learning practical, in Advances in kernel methods: support vector learning. 1999, MIT press. p. 169.

[22] Guo, G., et al. KNN model-based approach in classification. in OTM Confederated International Conferences" On the Move to Meaningful Internet Systems". 2003. Springer.

[23] Banfield, R.E., et al. A comparison of ensemble creation techniques. in International Workshop on Multiple Classifier Systems. 2004. Springer.

[24] Cramer, G., Ford, R., Hall, R., Estimation of toxic hazard—a decision tree approach. Food and cosmetics toxicology, 1976. 16(3): p. 255-276.

[25] Klecka, W.R., Iversen, G.R., Klecka, W.R., Discriminant analysis. Vol. 19. 1980: Sage.

[26] Lewis, D.D. Naive (Bayes) at forty: The independence assumption in information retrieval. in European conference on machine learning. 1998. Springer.

[27] Kleinbaum, D.G., et al., Logistic regression. 2002: Springer.

[28] Yildirim, M., Automatic classification and diagnosis of heart valve diseases using heart sounds with MFCC and proposed deep model. Concurrency and Computation: Practice and Experience, 2022: p. e7232.

[29] Tümen, V., SpiCoNET: A Hybrid Deep Learning Model to Diagnose COVID-19 and Pneumonia Using Chest X-Ray Images. Traitement du Signal, 2022. 39(4).

[30] Özbay, E., An active deep learning method for diabetic retinopathy detection in segmented fundus images using artificial bee colony algorithm. Artificial Intelligence Review, 2022. 1-28.

[31] Eroglu O, Keles E., Karlidag T, Kaygusuz I, Turker C, Yalcin S., Review of Our Tonsillectomy Indications Firat Med J, 2018. 23 (4): 178-183.

[32] Mazur, E., et al., Concurrent peritonsillar abscess and poststreptococcal reactive arthritis complicating acute streptococcal tonsillitis in a young healthy adult: a case report. BMC Infectious Diseases, 2015. 15(1): p. 1-5.