A Survey Analysis on Dental Caries Detection from RVG Images Using Deep Learning

P. Nageswari, Piyush Kumar Pareek, A. Suresh Kumar, Pai H. Aditya, M. S. Guru Prasad, and Manivel Kandasamy

1 Introduction

In the medical field, computer-supported evaluation software has been utilized to obtain second opinions. Deep learning techniques are difficult to use, but they have been incorporated with encouraging results for a number of health applications. With teeth, mouth is vital component of body, and each person's smile is unique. Poor oral health care leads to gum disease, dental caries, tooth loss, bone loss, and other dental issues. Oral illnesses impact ~4 billion people around the world (Singh et al. 2022). A developing area of study in the healthcare industry is medical imaging. Medical imaging is crucial for the early identification, diagnosis, and treatment of disorders. Dental caries is one of the most common conditions affecting people of all ages globally. Using radiovisiography (RVG) or dental X-ray images (Tomer et al. 2020), it could be challenging to spot dental caries in its early stages. Tooth decay,

P. Nageswari
Department of Mathematics, Mahendra Institute of Technology, Namakkal, Tamil Nadu 637503, India
e-mail: nageswarip@mahendratech.org

P. K. Pareek
Computer Science and Engineering, NITTE Meenakshi Institute of Technology, Bangalore 560064, India

A. Suresh Kumar (✉) · P. H. Aditya · M. S. Guru Prasad
School of Computer Science and Engineering, JAIN (Deemed-to-be University), Bengaluru, Karnataka 562112, India
e-mail: falsesuresh79.slm@gmail.com

M. S. Guru Prasad
e-mail: guruprasad.cse@geu.ac.in

M. Kandasamy
Unitedworld School of Computational Intelligence, Karnavati University, Gandhinagar, Gujarat 382422, India

N. R. Shetty et al. (eds.), *Advances in Communication and Applications*, Lecture Notes in Electrical Engineering 1105, https://doi.org/10.1007/978-981-99-7633-1_12

often known as dental caries, is a widespread chronic ailment that affects people all over the world and is caused by the erosion of dental enamel. The bacteria that cause dental decay produce acid that damages tooth enamel, eventually causing a small hole in the tooth. Formation of dental caries is though multifarious contact among acid producing bacteria and fermentable carbohydrate, teeth, saliva (Arora et al. 2020). Digital radiography has been used in dentistry for the past 20 years, and dental professionals are using it more frequently every day. Dental radiography displays tooth decay and bone loss. The presence of a hidden dental structure is not visible during a physical examination (Singh et al 2021).

Deep learning has had a substantial impact on the medical industry for a number of years, yielding results that are well acknowledged (Chandrappa et al. 2021). Despite the database's small size, dental surveys show that several studies are carried out to detect and cure dental caries. There is no publicly accessible or industry-standard database that contains dental X-ray pictures (Naveen Kumar et al. 2019). A database of 1336 dental X-ray images were produced by the authors (Mohd et al. 2021; Rubidha et al. 2022; Gupta et al. 2022). Since dental caries can range in size from tiny to large, it can be difficult to characterize patterns using dental X-ray images for caries identification (Nisha Chandran and Gangodkar 2021). In order to accurately diagnose dental cavities from 2D dental radiographs, CNN, a deep learning architecture, must be used. The difficult task of image segmentation and feature extraction from dental X-ray pictures is characterizing the area of a tooth with caries (Haseena 2022; Rajesh et al. 2020).

In light of the importance of dental caries crowd detection and analysis, numerous reports are recommended (Jawahar et al. 2022). These articles frequently discuss computer vision-based CNN models to detect dental caries. In this study, the three methodologies are density-based, regression-based, and detection-based will be contrasted (Kirubakaran et al. 2021).

2 Methodology

Convolution neural network is designed to identify between dental caries photographs with a 2D shape and performs better in this area than other deep learning models (Balamurugan et al. 2022). The ability of CNN to extract features is quite good. CNN maintains the spatial interaction of pixels, in contrast to vector image processing techniques. We have gathered a collection of roughly 1336 radiovisiography (RVG) X-ray pictures from dental clinic in India, since the typical dental X-ray database is not online (Balamurugan et al. 2022). X-rays of the patient's mouth are taken using dental imaging sensors, and RVG images are stored using Kodak's Care Stream 6100 RVG software. The dimensions of each RVG image are 748×512. To convert .rvg files into the .jpg format, which is supported by practically all dental clinics, one needs a license copy of the RVG software (Parasuraman et al. 2023). From the X-ray image, each tooth was clipped independently. These teeth sample images have been cropped and shrunk to 224 × 224. Every image is reduced in size and saved in a folder labeled "caries" or "no caries."

Database size is increased by applying image augmentation methods like rotate image positive by 15°, rotate image negative by 15°, rotate image positive by 60°, rotate image negative by 60°, rotate image positive by 90°, rotate image negative by 90°, vertical flip, horizontal flip, and mixture of horizontal, vertical, and rotating images by angle 15. A sample dental database with teeth with and without caries is shown in Fig. 1a, b.

A. *VGG-16 Architecture*

VGG-16 would be a convolutional neural network plan that won the ILSVR challenge in 2014. This is now recognized as among the top optical prototype plans on the market. The much more notable feature of VGG-16 would be that, rather than expanding the number of hyper-limits, researchers focused upon creating 3 × 3 channel convolution layers with stage 1 using similar padding and max pool layer as stage 2's 2 × 2 channel. Throughout the architecture, it continually follows the convolutional and maximum pool layers' game plan (Fig. 2).

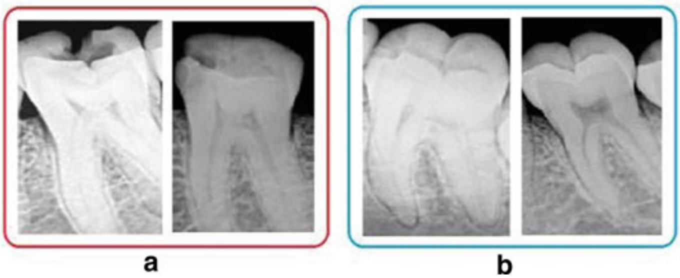


Fig. 1 Sample RVG images **a** caries, **b** no caries

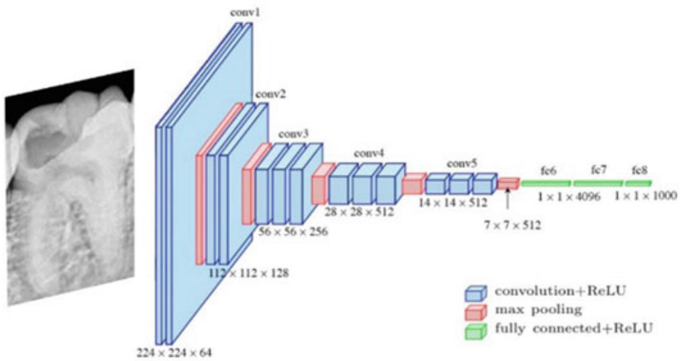


Fig. 2 VGG-16 architecture

Image Data Generator from the Keras. Preprocessing library to do VGG-16 model on dental X-ray images. The goal of Image Data Generator is to input information with names into the model as accurately as possible. It is a crucial class since it contains multiple capabilities such as rescaling, turning, zooming, flipping, and so on. The best part about this class is that it has no bearing on the information on the plate. This class hustles information while passing it on to the show. As a result, Image Data Generator will call all of the information inside the caries coordinator caries and the no-caries envelope no-caries. As a result, data can be successfully transmitted to a neural network.

The VGG-16 was designed using 1104 dental X-ray pictures of size 224 × 224 and tested on 121 images for 37 emphases, achieving a 75.3% testing accuracy. After the 38th emphases, the model exhibition begins to disintegrate.

B. *Inception V3*

A starting network is a deep neural network with a construction plan that includes repeated components known as beginning modules. It used a unique set of techniques to push execution, both in terms of speed and precision. Its continuous progress has resulted in the creation of a network with few variants.

(i) Inception v1

The size of striking parts in an image can range from exceedingly small to incredibly large. Because there can be such a wide range of information, choosing the proper piece size for convolution movement can be difficult. The more obtrusive piece is chosen for knowledge that is passed on all more locally, and the bigger portion is liked for information that is distributed all more throughout the world. Overfitting is a risk with really deep networks. It is also difficult to depart because the trend has resurfaced across the entire network. Stacking massive convolution exercises in a haphazard manner is computationally expensive. If channels of varying sizes had the choice to chip away at a comparative level, the network would essentially get nothing "more widespread" rather than "deeper." What do you think? demonstrates a "naive" conception module. It conducts convolution on commitment with three different channel sizes (1 × 1, 3 × 3, 5 × 5). Furthermore, maximum pooling has been completed. The results are connected and forwarded to the next module of the start-up process. Figure 3 shows the Naïve version of inception module.

To make it less expensive, deep neural networks are computationally expensive. Before the 3 × 3 and 5 × 5 convolutions, a 1 × 1 convolution layer is added. Although it may appear counterintuitive to add more activity, 1 × 1 convolutions are undeniably less expensive than 5 × 5 convolutions, and a lower number of information channels also help. Instead, then being shown before the max pooling layer, 1 × 1 convolution is shown after it.

(ii) Inception v2

Reduce the size of the illustrative bottleneck. When convolutions do not change components of information significantly, neural networks operate well, according to

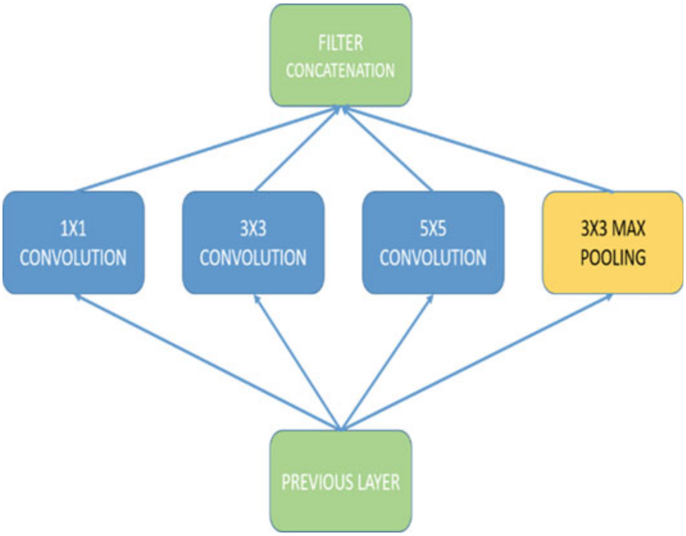
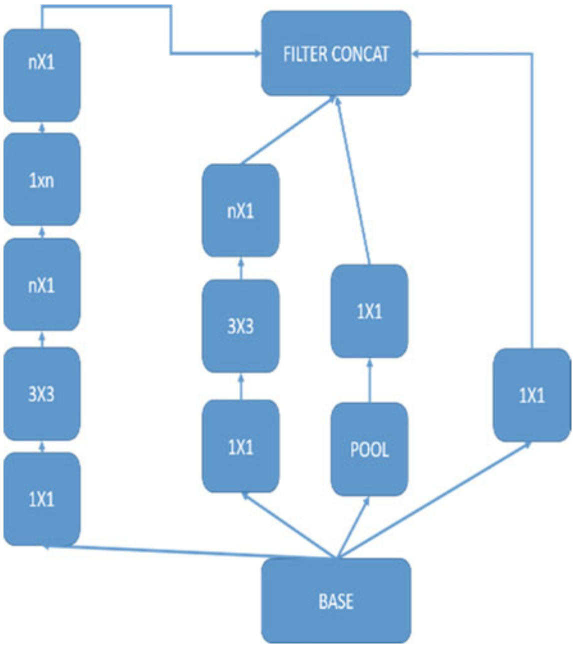


Fig. 3 Inception module Naïve version

Fig. 4 Stacking method causes lift in execution



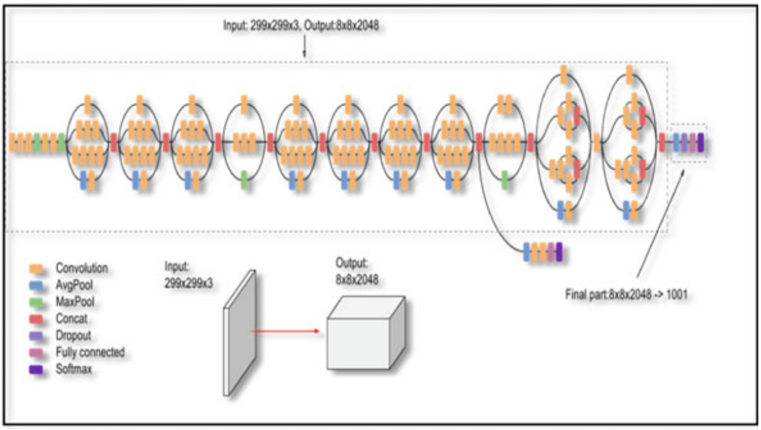


Fig. 5 Stacking method causes lift in execution

instinct. Excessive usage of lessening characteristics can result in data loss, which is regarded as "an example of a bottleneck" (Fig. 4).

(iii) *Inception v3*

Until near the end of the training procedure, when correctness was approaching immersion, auxiliary designations did not make much of a difference. It was decided to look into possible results to improve on Beginning v2 without necessarily modifying modules. As a result, version 3 was released, which included the RMS Prop Streamlining agent, factorized 7×7 convolutions, batch norm in auxiliary distributions, and name smoothing overfitting is avoided (Fig. 5).

C. *AlexNet*

AlexNet engineering consists of five convolutional layers and three completely associated layers, as seen in Fig. 6. A number of Convolutional Portions (channels) extract fascinating visual highlights. There are usually several parts of identical size in a single convolutional layer. AlexNet's primary Conv Layer, for example, has 96 bits of $11 \times 11 \times 3$ size. A section's width and height are usually equivalent, and its depth is the same for the range of channels. Following the initial two convolutional layers are Covering Max Pooling layers, which are shown next. The third, fourth, and fifth convolutional layers are directly linked. The Covering Max Pooling layer follows the fifth convolutional layer, and the consequence is the advancement of two entirely associated layers. Softmax categorization is handled by a second entirely related layer with 1000 class marks. After all convolution and fully related layers, ReLU nonlinearity is applied. Prior to doing pooling, ReLU nonlinearity of the first and second convolution layers is followed by a neighborhood standardization venture.

AlexNet's presentation will be ruined if any of the convolutional layers are removed. Engineering was done on dental X-ray pictures for 67 different ages, with a precision of 53.98%.

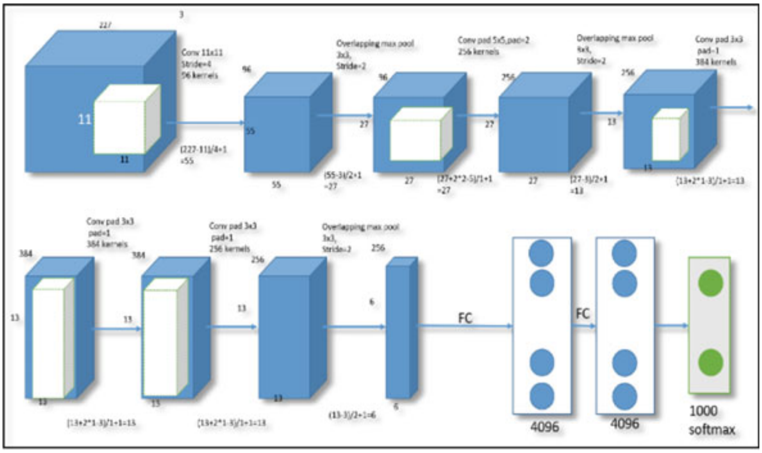


Fig. 6 Stacking method causes lift in execution

3 Proposed Architecture

Proposed architecture of CNN for dental caries detection (CCNN DCD) is given in Fig. 7. Convolutional neural networks, often known as CNNs or ConvNets, are a type of neural network that spends a lot of time processing data with lattice-like geography, such as images. CNN automatically distinguishes relevant elements with little to no human intervention. CNNs, like traditional neural networks, were energized by neurons in human and animal brains. Three important benefits of CNN were found by Singh et al (2021): similar representations, insufficient connections, and parameter allocation. In contrast to conventional fully associated (FC) networks, CNN uses shared loads and neighborhood linkages to use 2D information structures such as picture signals. This activity uses a small number of borders, which simplifies both preparation interaction and network velocities. A typical type of CNN, known as a multilayer perceptron (MLP), consists of several convolution layers preceding sub-testing (pooling) levels, with FC layers serving as completion layers. The advantages of employing CNNs over other classic neural networks in computer vision are almost impossible to ignore. A multilayer perceptron is a form of CNN that has numerous convolution layers before the sub-testing (pooling) levels, with FC layers functioning as completion layers.

4 Results and Discussion

In order to test and trained 1300 images out of 6000 RVG scanned images received from medical center in India. Out of those, 1300 images are scrutinized for training, and 100 images for verification and validation to access the performance of a system.

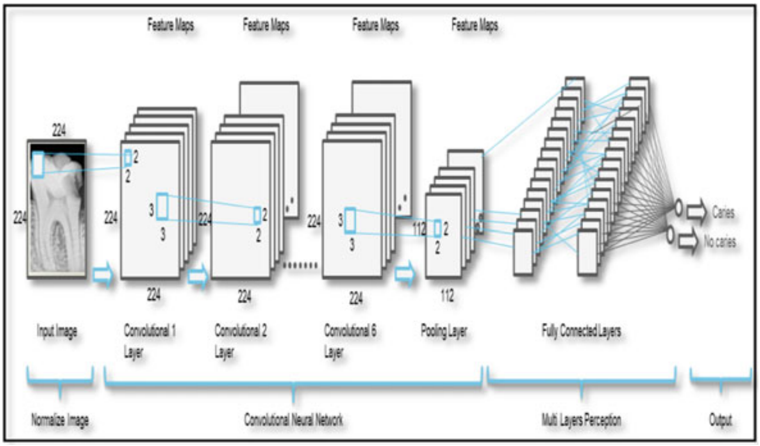


Fig. 7 CCNN architecture

This system is developed using Keras Python and TensorFlow libraries with the help of trained with Google Colab. The comparative analysis has been trained with Inception3, AlexNet, VGG-16, and proposed model.

Table 1 its showing compares the study of deep learning models with the suggested CCNN DCD model. Our model has accurately categorized 100 test photos. As a result, we were able to attain an acceptable accuracy of 93.2% achieved using given Eqs. 1, 2 and it is shown in Fig. 8.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} * 100, \tag{1}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} * 100. \tag{2}$$

Figure 9 shows the testing and trained accuracies of proposed system model with 100 epochs.

Table 1 Comparative study with CCNN

DL model	No. of epochs	Practice database size	Testing database size	Accuracy (%)
Inception3	80	1300	100	87.2
AlexNet	69	1300	100	61.89
VGG-16	37	1300	100	78.6
CCNN—proposed model	100	1300	100	93.8

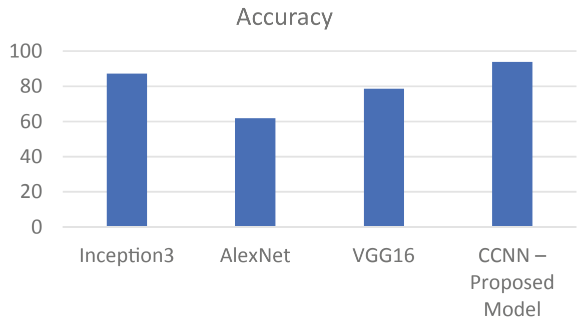


Fig. 8 CCNN for accuracy comparatives

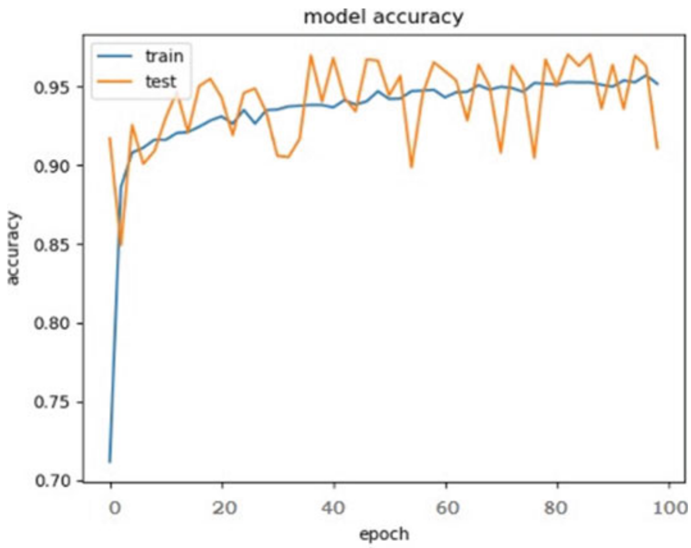


Fig. 9 CCNN for test and train accuracy models

5 Conclusion

In this paper, RVG pictures are classified using a specific, unique customized convolution network to detect for dental issues. In order to increase datasets and improve the image's visual readability, this section defines picture improvement approaches with a number of enhancement processes that highlight distinct differences between the image's attributes. The image is then double thresholder and used as a framework example to designate the focuses around the jaw. The unnecessary component was then removed after separating the upper and lower jaws in order to separate the individual teeth. A CCNN computation for dental disease distinguishing proof was employed to successfully detect impacted caries, abscesses, impacted third molars, and root fragments with the accuracy rate of 93.8% has been achieved.

References

Arora YK et al (2020) A Novel approach for detection and classification of rice leaf diseases. In: Springer—Advance in Intelligent systems and computing, vol 1155. https://doi.org/10.1007/978-981-15-4029-5_10

Balamurugan M, Kumaresan M, Haripriya V, Annamalai S, Bhuvana J (2022) Secured cloud computing for medical database monitoring using machine learning techniques. In: JeenaJacob I, Selvanayaki KS, Bestak R (eds) Notes in networks and systems, vol 444. Springer, Singapore, pp 91–109

Chandrappa S et al (2021) A novel approach to glaucoma screening using optic nerve head through image fusion and fractal geometry. J Math Comput Sci 11:4284–4300

Gupta AD et al. (2022) A framework of internet of things (IoT) for the manufacturing and image classification system. In: International conference on advance computing and innovative technologies in engineering (ICACITE), pp 293–297

Haseena S et al (2022) (2022) Prediction of the age and gen and gender based on human face images based on deep learning algorithm. Comput Math Methods Med 2022:1413597

Jawahar M, Anbarasi LJ, Ravi V, Prassanna J, Graceline Jasmine S, Manikandan R, Sekaran R, Kannan S (2022) CovMnet–deep learning model for classifying coronavirus (COVID-19). Health Technol 12:1009–1024

Kirubakaran J, Venkatesan GKDP, Sampath Kumar K, Kumaresan M, Annamalai S (2021) Echo state learned compositional pattern neural networks for the early diagnosis of cancer on the internet of medical things platform. J Ambient Intell Human Comput 12(3):3303–3316

Mohd N et al (2021) A survey: heart disease prediction using machine learning techniques. Webology 18(4):441–452

Naveen Kumar HN et al (2019) Optimization in feature extraction schemes on static images to improve the performance of automatic facial expression recognition systems. Int J Comput Sci Eng 7(6):1104–1109

Nisha Chandran S, Gangodkar D (2021) A novel image retrieval technique based on semi supervised clustering. Multimed Tools Appl 80:35741–35769

Parasuraman M, Munnangi AK, Rajeyyagari S, Sekaran R, Ramachandran M (2023) 5G wireless network-based cybersecurity analysis using software defined Phy_HetNets and Boltzmann encoder convolutional basis neural network. In: Hassanien AE, Castillo O, Anand S, Jaiswal A (eds) International Conference on Innovative Computing and Communications. ICICC 2023. Lecture Notes in Networks and Systems, vol 703. Springer, Singapore. https://doi.org/10.1007/978-981-99-3315-0_10

Rajesh E et al (2020) Analysis and Prediction of Liver Disease using Deep Learning. Int J Adv Sci Technol 29(11):2271–2278

Rubidha Devi D et al (2022) Enhancing orthopedic surgery and treatment using artificial intelligence and its application in health and dietary welfare. Hindawi J Food Qual 2022:7734650. https://doi.org/10.1155/2022/7734650

Singh AS et al (2021) Colour detection in objects using NIN implemented CNN. Special Ugdymas 1(43):3989–4001

Singh S et al (2022) Automated deep learning based disease prediction using skin health recored, issues, challenges and future directions. In: International conference on Electronics and renewable systems, pp 638–643. https://doi.org/10.1109/ICEARS53579.2022.9752422

Tomer V et al (2020) A selection method for computing the ensemble size of base classfier in multiple classifier system. In: Applied vision and Image processing. Springer—Advance in Intelligent systems and computing, vol 1155. https://doi.org/10.1007/978-981-15-4029-5_23