Detection and Classification of Age-Related Macular Degeneration Using Integration of DenseNet169 and Convolutional Neural Network

F. Ajesh[1,2(✉)] ⓘ and Ajith Abraham[1,3]

[1] Machine Intelligence Research Labs (MIR Labs) Scientific Network for Innovation and Research Excellence, Auburn, WA 98071, USA
`ajith.abraham@ieee.org`
[2] Department of Computer Science and Engineering, Sree Buddha College of Engineering, Alappuzha, Kerala, India
[3] Center for Artificial Intelligence, Innopolis University, Innopolis, Russia

Abstract. Age-related macular degeneration (AMD) is a leading cause of vision loss and blindness around the world. With an increase in age, the number of people impacted by the disease is observed to be growing. Knowledge about the occurrence of AMD should be used to develop appropriate eye care for these people. So, in this paper, we present an AMD detection and classification using DenseNetCNN. Data is collected from various repositories such as AREDS, Optretina and STARE. These are initially pre-processed using the histogram equalization technique. Then it is passed to feature extraction technique where GLCM comes in hand for extracting required features and finally passed to quintessential process which is the classification where DenseNet169 + CNN comes in play for effective classification. We have evaluated our model under accuracy, sensitivity, specificity performance measure and is compared with other pre-trained models like VGG16, ResNet50, GoogleNet, MobileNet and Inception V3 in which our model outperforms other state-of-art models with 98% of accuracy.

Keywords: Age-related macular degeneration · Ant colony optimisation · Classification · Convolutional neural network

1 Introduction

In Western countries, age-related macular degeneration (AMD) is the primary reason for vision loss and irreversible blindness of the aged [1]. It encompasses a wide range of macula problems. Although the initial stage of AMD is silent, examination of the retina can reveal tiny lesions known as drusen. The appearance of hemorrhages (wet AMD) or the formation of regional atrophy are both signs of progression of the disease, as is a growth in the size or quantity of drusen (late dry AMD). A Study [2] developed a clinical categorization for AMD. It is divided into four categories: non-AMD, mild, moderate,

Fig. 1. a) Vision of normal people. b) Vision of person with AMD

and progressive AMD [3]. Figure 1 (a) depicts the vision of normal people and 1(b), the vision of a person with AMD.

SD-OCT (Spectral-Domain Optical Coherence Tomography) is widely used equipment that explains specific AMD findings such as drusen, intra-retinal fluid (IRF), sub-retinal fluid (SRF), sub-retinal hyper-reflective substance, which includes hemorrhage and retinal colourant epithelium detachment, among others [12]. Exudative modifications (intra- and sub-retinal fluid and hemorrhage) are the most common reasons for many doctors to start anti-VEGF medication and monitor its effectiveness [13]. For the big clinical experiment discussed earlier [14], zero tolerance was used. As a result, the volume of OCT data that needs to be analyzed is growing faster than clinical capacity [15].

Machine-learning advances have found an answer for interpreting vast volumes of medical picture data resulting from repeated patient treatment and follow-up observations [16]. Convolutional neural networks (CNNs) have made tremendous progress in the processing of clinical images [18]. CNNs have already been used in ophthalmology for the categorization of diabetic retinopathy on fundus images, visual field evaluation of patients, and assessment of paediatric nuclear cataracts, among other applications [19]. Computerized detection of AMD characteristics in OCT and fundus pictures, anti-VEGF drug counselling, and disease progression tracking have all been done with neural networks [19]. Although several researchers have categorized the exudative element of AMD for computerized segmentation [21], we are unaware of any publications that categorize exudative alterations with AMD in deep learning models excluding segmentation.

This paper focuses on an accurate and efficient model for detection and classification of AMD in which key points are as mentioned below:

1. The detection of AMD by Deep Learning
2. Datasets used are AREDS, Optretina and STARE.
3. DenseNet169 + CNN is used for Classification.
4. An accuracy of 98% is achieved.

The remaining sections organized as follows: Sect. 2 depicts the works related to this paper that is been so far done by research while Sect. 3 illustrates the methodologies of the proposed model. Section 4 is focused on the simulation analysis of proposed model and finally the paper concludes with Sect. 5.

2 Literature Review

Using the Optretina database and CNN approach, Zapta et al. [17] used AI to detect retinal fundus images, quality verification, laterality analysis, macular degeneration, and potential glaucoma. They had a 94.7% accuracy rate. A Convolutional Neural Network is much slower due to procedures such as a max pool [4]. The training technique takes a long time if the computer does not have a powerful GPU and a ConvNet requires a huge database to analyze and train the neural network.

Motozawa et al. [26] used the STARE database and the CNN approach to develop OCT-based DL (Deep Learning) models for distinguishing normal and AMD, as well as exudative and non-exudative AMD alterations. They had a 96% accuracy rate. Computational power (depends on network design and data amount) and Sophisticated architecture are two drawbacks of their job (Not every time).

With the database AREDS and CNN, Keenan et al. [6] developed a DL system for automatic detection of geographic atrophy from color fundus pictures. They achieved an accuracy of 92.7%. The disadvantage of their work is that it does not work well with high dimensions and requires more computational resources.

Grassman [20] used the CORRA database and the RNN approach to construct a DL system to predict Age-Related Eye Disease (AREDS) research intensity scale for AMD using color fundus photography. They had an accuracy percentage of 92.7%. The drawbacks of their work include the fact that neural networks are black boxes, meaning we don't know whether each independent variable influences the dependent variables. Traditional CPU training is both technologically incompetent and time-intensive, and neural network training data is crucial [23]. As a result, over-fitting and generalization become a worry. The method is more dependent on training data and can be customized to meet your specific requirements.

3 Methodology

Here the author tries to propose a new methodology for the detection and classification of AMD using a hybrid structure of Densenet 169 and CNN [5]. Figure 2 depicts the diagram of the methodology of the proposed model. Here, initially, we collect datasets from repositories like AREDS, Optretina and STARE which is followed by proposing this dataset as it contains noises and other anomalies. For that, we use histogram equalization processing technique. Further, extracting features using GLCM is done. Finally, passes to the main stage which is classification [25]. Here we use the DENSENET169 neural network for accurately classifying images into disease/no-disease or dry/wet ARMD.
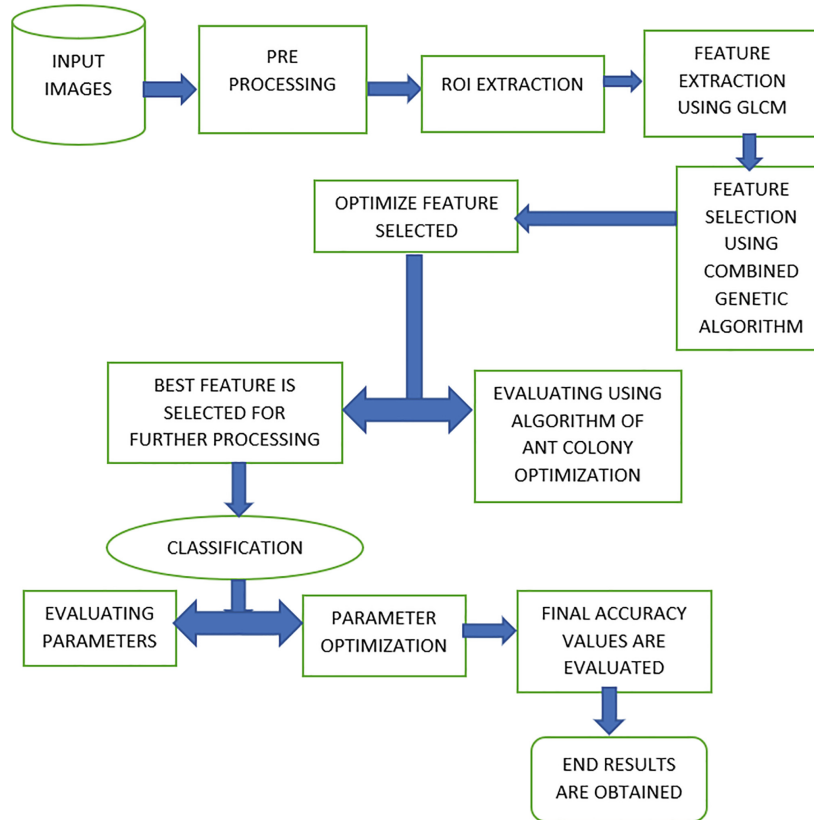
Fig. 2. Methodology of the proposed model

3.1 Dataset

The repositories used are AREDS, Optretina and STARE.

- The AREDS is a multi-centre, panel study of AMD and cataracts. The AREDS study included a clinical investigation of large quantity mineral and vitamin supplements for AMD [7] and a clinical investigation of large quantity vitamin supplements for cataracts, as well as natural history data. Participants in the AREDS study had to be between the ages of 55 and 80 at the time of enrollment, and they had to be in healthy conditions which would otherwise make a longitudinal observation or drug administration difficult or impossible. 4,757 participants were divided into four AMD categories based on fundus images assessed by a centralised reading centre, finest visual acuity, and ophthalmologic examinations.
- Optretina has been doing telemedicine screenings at optical centres since 2013, and in workplace offices and private firms since 2017. The Optretina tagged contents, which contain 306,302 retinal pictures, were used in this study. NMC (colour fundus and red-free) and optical coherence tomography are used to create the images (OCT) [8]. Images from various types and brands of cameras were included in the dataset.
- The STARE dataset (Structured Analysis of the Retina) is a retinal vascular segmentation dataset. It includes 20 colour fundus photos that are all the same size (700605).

3.2 Pre-processing

Once the data is collected, preprocessing should be done to remove any noises or anomalies. For the cleansing purpose, we used histogram equalization. Histogram equalization is a contrast correction approach in image processing that uses the image's histogram. For that, we need to convert the given image into a greyscale version. In greyscale conversion, we normalize the non-uniformities and then enhances the contrast of the image. Then this converted image is given for histogram equalization process where it enhances image based on the intensity values and thereby receiving a pre-processed output which will then be passed on for extraction process [9].

3.3 Feature Extraction

In the proposed model, the GLCM technique is used to derive information from the image. The proportion of co-occurring values at a given interval is represented by a grey level co-occurrence matrix (GLCM), which is a matrix formed over an image [22]. In a GLCM matrix, the quantity of grey levels, G, in an image equals the number of rows and columns. As a result, using statistical characteristics is one of the earliest methodologies proposed in the image analysis literature. Haralick [14] recommended using a co-occurrence matrix. It evaluates the connection between the two adjacent pixels, the first of which is referred to as a reference pixel and the second as a neighbour pixel. It's also known as a grey tone spatial dependence matrix [10]. It is used to identify the texture of an image by tabulating the intensity values of an image and the frequency with which they appear. The five characteristics we extracted are homogeneity, correlation, co-occurrence, energy, and entropy. There are two-pixel values included in GLCM, standard pixel and neighbor pixel. The neighbor pixel is picked as the pixel to the right of each reference pixel. The GLCM matrix is a square matrix with Ni as the no: of grey levels. To get element [j,k], divide the total number of such comparisons by the number of times the pixel value of j is next to the pixel value of k. The features extracted from the GLCM are contrast, correlation, energy, homogeneity and entropy. Table 1 depicts the corresponding equations of the features extracted.

3.4 Feature Selection and Feature Optimization

Feature selection is a technique for choosing a collection of extracted features or creating factors with the greatest classification results. This method prevents model overfitting by removing potentially irrelevant or misleading information. In other terms, it identifies key characteristics which can be utilized to distinguish healthy from unhealthy images. Ant colony optimization (ACO) is a successful approach that is useful in the solution of NP-hard combination optimization problems and it has been widely applied in GWAS. The basic idea of ACO is to express the feasible solutions of optimization problems with ant paths and use overall paths of the ant group to constitute the solution space of optimization problems. Ants on relatively short paths tend to release more pheromone. With the passage of time, pheromone concentration that accumulates on the short paths gradually increases and more and more ants choose the paths. Eventually, all the ants

Table 1. Statistical features of the features used in feature extraction

Extracted features	Equation
Mean	$M = \left(\frac{1}{m \times n}\right) \sum\limits_{x=0}^{m=1} \sum\limits_{y=0}^{n=1} f(x, y)$
Standard deviation	$SD(\sigma) = \sqrt{\left(\frac{1}{m \times n}\right) \sum\limits_{x=0}^{m=1} \sum\limits_{y=0}^{n=1} (f(x, y) - M)^2}$
Entropy	$E = - \sum\limits_{x=0}^{m=1} \sum\limits_{y=0}^{n=1} f(x, y)^2 f(x, y)^2$
Skewness	$S_k(X) = \left(\frac{1}{mn}\right) \frac{\sum (f(x,y) - M)^3}{SD^3}$
Kurtosis	$K_{urt}(x) = \left(\frac{1}{mn}\right) \frac{\sum (f(x,y) - M)^4}{SD^4}$
Contrast	$C_{on} = \sum_{k=0}^{m-1} \sum_{y=0}^{n-1} (x - y)^2 f(x + y)$
Correlation	$Corr = \frac{\sum_{x=0}^{m-1} \sum_{y=0}^{n-1} (x,y) f(x,y) - M_x M_y}{\sigma_x \sigma_y}$
Coarseness	$C_{ness} = \frac{1}{2^{m+n}} \sum_{k=0}^{m-1} \sum_{y=0}^{n-1} f(x, y)$

will gather on the optimal path under positive feedback, which exactly corresponds to the optimal solution of the optimization problem.

In this study, a genetic optimization strategy formed using crossover and mutation operators in the genetic algorithm was combined with an ant colony optimization shortest-path method to achieve feature selection. Holland created the Genetic Algorithm (GA), which is a method of addressing an optimization issue that operates similarly to biological evolution. In a genetic algorithm, a starting population of solutions (similar to chromosomes) is selected and subjected to iterative change. Every member in the current population has a fitness value assigned to them. The fitness value is determined by training the prediction model with the training data set and then calculating the selection error. The lower the value of selection error, the lower the fitness. As a result, those with a higher fitness value will be chosen to create the next population. For a predetermined number of generations, the algorithm continues until the best answer throughout the evolution process does not shift to a better value. This predefined value can be 20% or 30% of the generation number determined by the best solution so far.

For example, at generation 50, the algorithm achieves a value of 200, which does not change for 15 generations (30% of 50), causing the process to stop. Our primary objective was to distinguish between dry and moist ARMD [11], as well as ARMD without symptoms. To obtain a P-value, variance analysis was used to choose the top-ranking energy and entropy parameters. For categorization, the top ten statistically significant (P 0.05) variables (1 energy, 3 entropy, 6 other nonlinear) were chosen. The optimization is
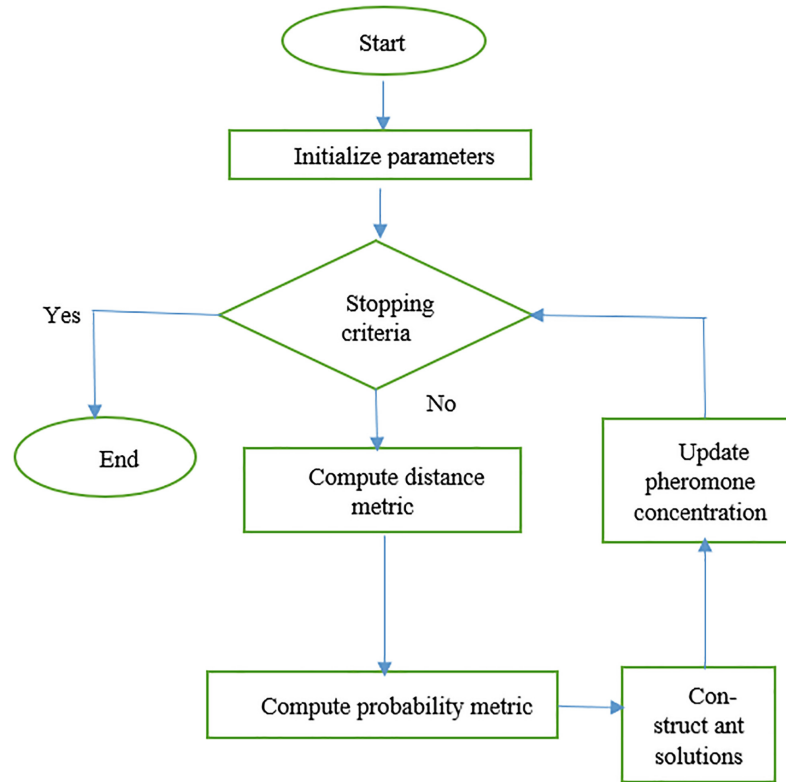
Fig. 3. Flowchart of ant colony optimization

done by Ant Colony Optimization. The flowchart of Ant Colony Optimization is depicted in Fig. 3.

3.5 Classification

This is the most crucial process of the entire system in which we use several ML or DL techniques to predict the desired output. So, in this paper, for diagnosing the AMD anddisease/no-disease or dry/wet ARMD, we use DenseNet169 + CNN. It is a Fully Convolutional Neural Network (FCN) that is a simultaneously trained learning network with convolutional filters in place of fully linked layers as judgement layers. By linking the output neurons of wholly attached levels to all input neurons, this alteration on the top level's aids in the reduction of data connected to a place caused by fully linked areas. DenseNet169 + CNN's defining feature is its ability to reuse features extracted and boost characteristic dispersion by establishing a straight link among each layer and every subsequent layer. Dense, transition down, and transition up are the three essential blocks of the DenseNet169 + CNN network. A batch normalization step follows a ReLU as an input signal, a 3×3 convolution layer, and a dropout layer with a 0.2 decreasing rate in the Dense block (DB). A batch normalization layer, a ReLU as an input signal, a 3×3 convolution layer, a dropout layer with a 0.2 dropping rate, and a 2×2 Max pooling layer make up a transition down (TD) block. Three transposed convolution layers make up a transition Up (TU) block. It's worth noting that batch normalization and dropout

can both help to reduce overfitting. The network may need one, both, or none of these, based on the extent of overfitting. We discovered that integrating both dropout and batch normalization in our network improves performance on this issue. The network's architecture includes a 3×3 convolution layer on the input, five dense blocks with 4, 5, 7, 10, and 12 layers each, a transition down component, one dense block with 15 layers in the final layer of the down-sampling path (bottleneck), five transitions up blocks with dense blocks of 12, 10, 7, 5, and 4 layers, and a 1×1 convoluted layer followed by a non-linearity indicated by the Softmax function. RMSprop [24], a stochastic gradient descent optimization technique, is utilized for training the network with a rate of learning of 10–3 in 120 epochs and a 30 epoch early-stop condition. The images are enhanced with vertical flips and irregular cropping to artificially increase the number of images. Uniform distribution was used to initialize the network's weights, and the loss function was cross-entropy. Testing of stages can begin once the model has been developed by utilizing the trained model to section the images in the test set. Figure 4 depicts the architecture of DenseNet169 + CNN.
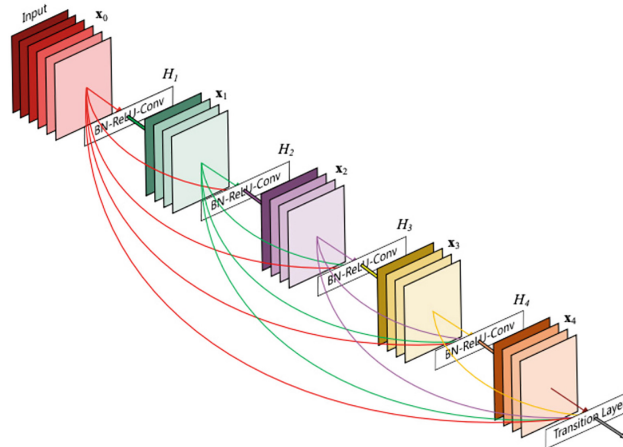
Fig. 4. Architecture of DenseNet169 + CNN

4   Experimental Results

The proposed DenseNet169 + CNN model is compared with the existing cut-edge models such as VGG16, ResNet50, GoogleNet, MobileNet and Inception V3 and their performance is measured using several measures such as Accuracy, Specificity, Sensitivity, Precision, F-measure and confusion matrix.
.

Table 2. Metrics of performance measures

| Metrics of performance measure	Mathematical description
1. Sensitivity, TPR	$= \frac{TP}{TP+FN}$
2. Specificity, S	$= \frac{TN}{FP+TN}$
3. Precision	$= \frac{TP}{TP+FP}$
4. Accuracy	$= \frac{TP+TN}{TP+FN+TP+TN}$
5. F Score	$= \frac{2TP}{2TP+FN+FP}$

Here, True positive (TP): You predicted a positive and it turned out to be correct; True Negative (TN): The negative you predicted is correct; False Positive (FP): You predicted something positive but it wasn't true, and False Negative (FN): You predicted something negative but it wasn't true. Table 2 shows the metrics of performance measures.

Table 3 illustrates the evaluation results of DenseNet169 + CNN with other models in which CNN gives better results. Figure 5 gives the graphical representation of the confusion matrix.

Table 3. Confusion matrix

| Classification models	TP	FN	FP	TN
VGG16	142	20	12	59
GoogleNet	136	28	21	43
InceptionV3	120	38	38	30
ResNet50	132	29	27	41
MobileNet	149	7	5	70
DenseNet169 + CNN (Proposed)	160	1	1	30

Table 4 represents the comparison of the proposed model with other models based on performance for images 1, 2 and 3 respectively. Figures 6, 7 and 8 depict a graph of overall performance analysis of images 1, 2 and 3. Table 5 depicts comparison of proposed model with state-of-art works.
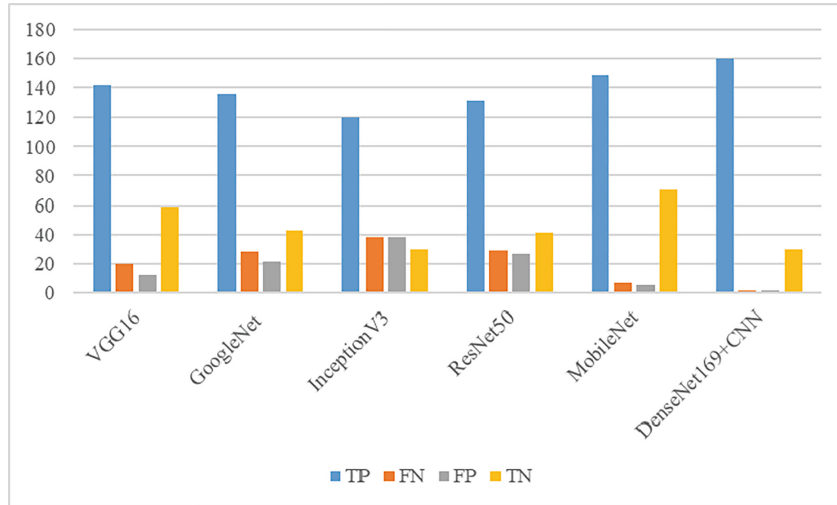
Fig. 5. Graphical representation of confusion matrix

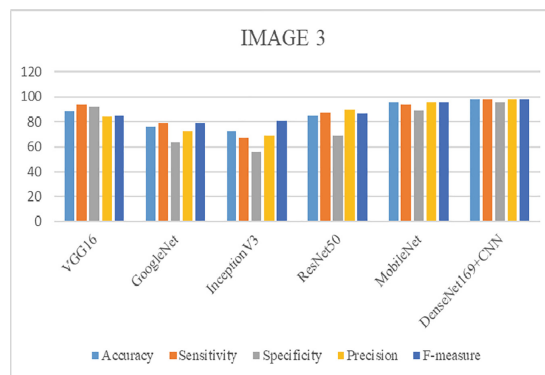**Table 4.** Performance analysis of images 1,2 and 3.

	Metrics	VGG16	GoogleNet	inceptionV3	ResNet50	MobileNet	Densenet169 + CNN (Proposed)
Image 1	Accuracy	91	75.1	69	82	95	98
	Sensitivity	91.78	81.59	78.08	86.6	95.77	98.5
	Specificity	88.8	60	44.44	68	93.33	96.5
	Precision	94.7	82.6	79.17	89.04	97.14	98.5
	F-measure	93.71	82.09	78.62	87.84	96.45	98
Image 2	Accuracy	90	79	64	74	93	98
	Sensitivity	84.1	89	78.01	88.11	93.42	97.95
	Specificity	88	64.21	54.24	69.21	94.11	96.5
	Precision	92.12	74.54	89.20	88.24	98	98.5
	F-measure	93.71	82.09	78.62	87.84	96.45	98
Image 3	Accuracy	89	76.21	72	85	96.02	98
	Sensitivity	94.21	78.59	67.08	87.56	93.70	98.15
	Specificity	92.12	64	55.54	69.12	89.12	96
	Precision	84.17	72.46	69.19	90.04	96.14	98
	F-measure	84.71	79.09	81	86.94	96	97.95
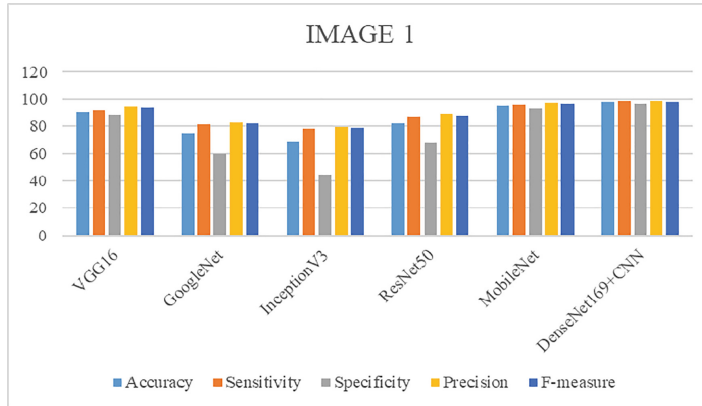
Fig. 6. Overall performance analysis of image 1
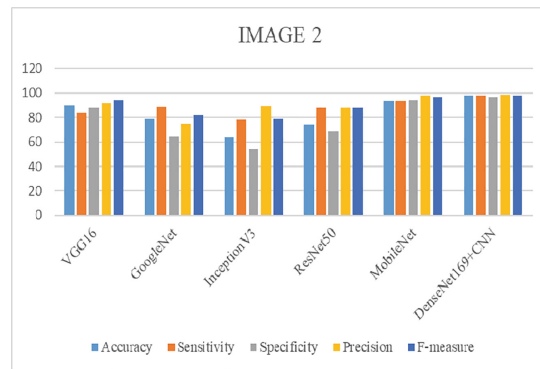
Fig. 7. Overall performance analysis of image 2

Fig. 8. Overall performance analysis of image 3

Table 5. Comparison of proposed model with state of art works

Author	Database	Method	Accuracy (%)
Zepta et al. [17]	Optretina	AI	94.7
Motozoa et al. [26]	Stare	CNN	96
Keenan et al. [6]	Areds	DL	92.7
Grassman et al. [20]	Areds	DL	92.7
Proposed model	Areds, Optretina and Stare	DenseNet169 + CNN	98

5   Conclusions

This research introduces a new DenseNet169 + CNN model for AMD detection and classification. We have suggested a new methodology with a few steps, which makes the model more cut edge one among the present ones. This methodology used the DenseNet169 + CNN more specific and sensitive compared to other models. Here, we used DenseNet169 + CNN for better evaluation and it clearly classified the given image into disease/non-disease or wet/dry AMD. This model is compared with other present models and we obtained a better result of 98% accuracy. Also, this paper is much useful for other researchers in helping them to bring other hybrid models for even better evaluation of AMD.

References

1. Stark, K., Olden, M., Brandl, C.: The german AugUR study: study protocol of a prospective study to investigate chronic diseases in the elderly. BMC Geriatr (2015)
2. Prenner, J.L., Halperin, L.S., Rycroft, C., Hogue, S., Williams Liu, Z., Seibert, R.: Disease burden in the treatment of age-related macular degeneration: findings from a time-and-motion study. Am. J. Ophthalmol. **160**(4), 725–731e1 (2015)
3. National Eye Institute: Facts about age-related macular degeneration. Available: https://nei.nih.gov/health/maculardegen/armd_facts (2015). Accessed 28 Jul 2017
4. Koh, J.E.W., Ng, E.Y.K., Bhandary, S.V., Laude, A., Acharya, U.R.: Automated detection of retinal health using PHOG and SURF features extracted from fundus images. Appl. Intell. **48**(5), 1379–1393 (2017). https://doi.org/10.1007/s10489-017-1048-3
5. Steinberg, J., Uibel, S., Berndt, T., Müller, D., Quarcoo, D., Groneberg, D.A.: Zentralblatt für Arbeitsmedizin, Arbeitsschutz und Ergonomie **61**(8), 270–286 (2011). https://doi.org/10.1007/BF03345002
6. Keenan, T.D., et al.: Progression of Geo- ´graphic atrophy in age-related macular degeneration: AREDS2. Ophthalmology **125**, 1913–1928 (2018)
7. Tan, J.H., Acharya, U.R., Bhandary, S.V., Chua, K.C., Sivaprasad, S.: Segmentation of optic disc, fovea and retinal vasculature using a single convolutional neural network. J. Comput. Sci. **20**, 70–79 (2017)

8. He, K., Zhang, X., Ren, S., Sun, J.: Delving deep into rectifiers: surpassing human-level performance on imagenet classification. In: Proceedings of the IEEE International Conference on Computer Vision, vol. 11–18–Dece, pp. 1026–1034 (2016)

9. Köse, C., Şevik, U., Gençalioğlu, O.: Automatic segmentation of age-related macular degeneration in retinal fundus images. Comput. Biol. Med. **38**(5), 611–619 (2008)

10. Köse, C., Şevik, U., Gençalioğlu, O., İkibaş, C., Kayıkçıoğlu, T.: A statistical segmentation method for measuring age-related macular degeneration in retinal fundus images. J. Med. Syst. **34**(1), 1–13 (2010)

11. Ferris, F.L.: Clinical classification of age-related macular degeneration. Ophthalmology **120**(4), 844–851 (2013)

12. Kuhn, M.: Building predictive models in R using the caret package. J Stat Softw. (2008)

13. Aiello, S., Eckstrand, E., Fu, A.: Fast scalable R with H20. In: Grün, B., et al. (eds.) Foundation for Open Access Statistics. ISSN (2015)

14. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet Classification with Deep Convolutional Neural Networks. Curran Associates Inc (2012)

15. Szegedy, C., Liu, W., Yangqing, J.: Going deeper with convolutions. In: 2015 IEEE Conf. Comput. Vis. Pattern Recognit. IEEE, Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society Press, ISSN: 1063–6919

16. Gulshan, V., Peng, L., Coram, M.: Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA (2016)

17. Zapata, M.A., Royo-Fibla, D., Font, O.: Artificial intelligence to identify retinal fundus images, quality validation, laterality evaluation, macular degeneration, and suspected glaucoma. Clin Ophthalmol. **14**, 419–429 (2020)

18. Szegedy, C., Vanhoucke, V., Ioffe, S.: Rethinking the Inception Architecture for Computer Vision. Computing Research Repository (CoRR). abs/1512.0. Available at: https://arxiv.org/corr/home (2015)

19. Fritsche, L.G., Igl, W., Bailey, J.N.C.: A large genome-wide association study of age-related macular degeneration highlights contributions of rare and common variants. Nat. Genet. (2016)

20. Grassmann, F., Mengelkamp, J., Brandl, C.: A deep learning algorithm for prediction of age-related eye disease study severity scale for age-related macular degeneration from color fundus photography. Ophthalmology **125**, 1410–1420 (2018)

21. Grassmann, F., Ach, T., Brandl, C.: What does genetics tell us about age-related macular degeneration?. Annu. Rev. Vis. Sci. (2015)

22. Swartz, R., Loewenstein, A.: Early detection of age-related macular degeneration, Int. J. Retin. Vitr. Article Number 20, December 01 2015

23. Swaroop, A., Branham, K.E., Chen, W., Abecasis, G.: Genetic susceptibility to age-related macular degeneration: a paradigm for dissecting complex disease traits. Hum. Mol. Genet. (2007)

24. Holz, F.G., Bindewald-Wittich, A., Fleckenstein, M.: Progression of geographic atrophy and impact of fundus autofluorescence patterns in age-related macular degeneration. Am. J. Ophthalmol. (2007)

25. Ratnapriya, R., Chew, E.Y.: Age-related macular degeneration-clinical review and genetics update. Clin. Genet. **84**, 160–166 (2013)

26. Motozawa, N., et al.: Optical coherence tomography-based deep-learning models for classifying normal and age-related macular degeneration and exudative and non-exudative age-related macular degeneration changes. Ophthalmol Therapy **8**(4), 527–539 (2019). https://doi.org/10.1007/s40123-019-00207-y