Received: 5 July 2021 | Accepted: 3 December 2021

DOI: 10.1002/alz.12565

Alzheimer's & Dementia®
THE JOURNAL OF THE ALZHEIMER'S ASSOCIATION

ALTERNATE FORMAT RESEARCH ARTICLE

Triploid genetic algorithm for convolutional neural network–based diagnosis of mild cognitive impairment

Harsh Bhasin | **R. K. Agrawal** | **for Alzheimer's Disease Neuroimaging Initiative**

School of Computer and Systems Sciences, Jawaharlal Nehru University, Delhi, India

Correspondence
Harsh Bhasin, School of Computer and Systems Sciences, Jawaharlal Nehru University, New Mehrauli Road, JNU Ring Rd, New Delhi, India.
E-mail: i_harsh_bhasin@yahoo.com

Abstract

The diagnosis of mild cognitive impairment (MCI), which is deemed a formative phase of dementia, may greatly assist clinicians in delaying its headway toward dementia. This article proposes a deep learning approach based on a triploid genetic algorithm, a proposed variant of genetic algorithms, for classifying MCI converts and non-converts using structural magnetic resonance imaging data. It also explores the effect of the choice of activation functions and that of the selection of hyper-parameters on the performance of the model. The proposed work is a step toward automated convolutional neural networks. The performance of the proposed method is measured in terms of accuracy and empirical studies exhibit the preeminence of our proposed method over the existing ones. The proposed model results in a maximum accuracy of 0.97961. Thus, it may contribute to the effective diagnosis of MCI and may prove important in clinical settings.

KEYWORDS
convolutional neural networks, deep learning, magnetic resonance imaging, mild cognitive impairment, triploid genetic algorithm

1 | NARRATIVE

Mild cognitive impairment (MCI) is considered the asymptomatic or preclinical stage in progressive cognitive decline leading to the more severe stage in dementia.[1,2] MCI is characterized by problems with memory, language, thinking, and judgment that are greater than normal age-related changes.[3,4] Symptoms of MCI may progress to Alzheimer's disease (AD) or another type of dementia, or remain stable for years, or even improve over time.[5] The annual conversion rate of MCI to AD is ≈15%.[6]

The challenge of accurate classification of MCI-converts (MCI-C) and MCI-non-converts (MCI-NC) is the central focus of this article; precision in the classification of asymptomatic people at potential risk is a vital need for future prevention studies or interventions to slow the progression of dementia and so apposite steps can be taken for slowing its headway toward dementia.

MCI can be detected by either clinical tests or brain scans. A medical professional can judge cognitive and behavioral changes and make a professional judgment about the presence or absence of MCI, by assessing the possible causes and the severity of the symptoms.[4] In recent years, numerous brain imaging modalities such as functional magnetic resonance imaging (fMRI),[7] positron emission tomography (PET),[8] structural-magnetic resonance imaging (s-MRI),[9–17] and so on have been used for diagnosing MCI. Brain-imaging studies show that the shrinkage of gray matter may be associated with MCI.[18]

Manual detection of MCI is not just time-consuming but may also be expensive and inconvenient. Many techniques have been proposed by the machine learning (ML) community to accomplish the above classification task. Conventional feature extraction methods have been successfully applied[19–23] to achieve this goal. These techniques have their advantages and downsides and may miss out on some discriminating patterns. Deep learning techniques parameterize multi-layer neural networks and learn the representation of data with multiple layers of abstraction.[24] The major difference between the conventional machine learning methods and deep learning algorithms is that the latter do not require feature extraction or selection, need little or no

image pre-processing, and generally result in a more objective classification. Convolutional neural networks (CNN) are architectures that encode certain properties of the input, assumed to be images, into the architecture.[25] This not only reduces the number of parameters to a great extent but also makes the networks efficient. The performance of the CNN model depends on the number of filters, choice of activation functions, and the number of units in the dense layer if the coarse structure of the network is known. The deep architectures perform better but are computationally expensive to train whereas shallow ones are computationally less expensive but may not perform aptly, if not crafted with due deliberation. While most of researchers decide the parameters of the CNN architectures using empirical analysis, automated networks have been used by some researchers.[26]

This work uses a proposed variant of the genetic algorithm (GA), called the triploid genetic algorithm (TGA), to find these hyper-parameters. The work also explores the applicability of a recently proposed periodic activation function for implicit neural representation.[27] This work also explored diploid genetic algorithms (DGA) for finding the hyper-parameters of a coarse CNN. To the best of our knowledge, no one has applied DGA for this task. The performance of the proposed method is judged against existing methods on publicly available Alzheimer's Disease Neuroimaging Initiative (ADNI) data. Experimental results suggest that the proposed model can extract relevant features to classify MCI-C and MCI-NC.

Empirical analysis suggests that TGA performs better than DGA. It was also observed that DGAs perform better compared to GAs for classifying MCI-C and MCI-NC volumes. Based on the results, it may be stated that activation functions affect the performance of the model. The proposed model shows superior performance vis-à-vis the conventional machine learning methods also.

Like other methods, the proposed model may not work for all datasets. In future work, we aim to explore ploidy further to achieve the goal of a fully automated CNN. Also, the performance of the proposed model depends on the crossover rate, the mutation rate, and the model is also computationally expensive. In addition to this, other unexplored activation functions may also improve the performance of the classification of MCI.

RESEARCH IN CONTEXT

1. **Systematic review**: The Classification of MCI has been done using both conventional Machine Learning methods, involving various feature extraction and feature selection techniques, and the Deep Learning methods. The research works that used Deep Learning, particularly CNN, can be further divided into two classes: those involving manually crafted CNN's[26] and those involving the automated CNN's. In the former, the hyper-parameters like the number of filters etc. are found manually using empirical analysis, whereas in the latter some search technique like Genetic Algorithm has been used to find these hyper-parameters. The proposed method uses TGA to find these hyperparameters. The method also takes into consideration the effect of activation functions on the performance of the model. Table 7 compares the performance of the proposed method, in terms of accuracy, with the state of the art. The proposed algorithm was implemented, and the experiments conducted brought forth some interesting points, discussed in the following subsection.
2. **Interpretation**: The contributions of this work are as follows:
    1. The development of a variant of Genetic Algorithm, called Triploid Genetic Algorithm, which performs better as compared to GA, and the three stated variants of DGA, for the classification of MCI.
    2. The development of a framework to find the hyper-parameters of a CNN, whose course structure is known. These hyper-parameters are selected on the basis of the given data.
    3. The work empirically proves that activation function plays an important part in the recital of a CNN. This work also explores the applicability of recently proposed SIREN activations.
3. **Future scope**: This work proposes a triploid to haploid conversion, for finding the hyper-parameters of CNN. In the future work, we aim to explore poly-ploidy, and carry out an extensive empirical analysis to find which, if any works well. To further this task we aim to use three different classification problems: MCI-C and MCI-NC, MCI and CN, and MCI-AD.

2 | CONSOLIDATED RESULTS AND STUDY DESIGN

Data used in the preparation of this article were obtained from the ADNI database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public–private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether MRI, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and to detect AD at an early stage.[28]

The ADNI database was queried for MCI-C and MCI-NC. This study uses 75 MCI-C and 112 MCI-NC processed Neuroimaging Informatics Technology Initiative images of patients. The protocol of data selection and image acquisition of the subjects has been adopted from Salvatore et al.[29] and takes into consideration age matching, required parameters, and so on. All patients had a Mini-Mental State Examination score between 18 and 27 and a Clinical Dementia Rating score of 0.5 or 1. The T1 weighted s-MRI images collected had the following parameters: TR = 3000 ms, TE = 3.6099 ms, and field strength = 1.5 Tesla. Table 1 shows the demographic information of the patients.

TABLE 1 Demographic information of the subjects

Group	MCI-NC (n = 112)	MCI-C (n = 75)
Female/ Male	51/61	33/42
Age (Mean ± SD)	74.83 ± 7.34	74.69 ± 7.28

Abbreviations: MCI-C, mild cognitive impairment converts; MCI-NC, mild cognitive impairment non-converts; SD, standard deviation

s-MRI has been generally used to identify MCI.[10–20] The process begins with pre-processing, for which statistical parametric mapping (SPM) is generally used.[30] The steps in the pre-processing of the s-MRI images include the following: (1) slice time correction, (2) head motion correction, (3) spatial normalization, (4) special smoothening, and (5) tissue segmentation, in which the brain tissues are separated into three tissue classes—namely gray matter, white matter, and cerebrospinal fluid. This work uses gray matter for building an ML-based model to diagnose MCI as its change has been found responsible for MCI in the literature.[18]

This work uses a CNN architecture that has three convolution layers. Each convolution layer is followed by an optional pooling layer. There are two fully connected layers, followed by a SoftMax. In the proposed model 30 slices from the top and 30 from the bottom were removed from each volume. The model extracts the feature vector of a given volume as follows. For each slice, the output of the penultimate layer is extracted. These are horizontally concatenated to form the feature vector of the volume. This is followed by feature selection using Fisher discriminant ratio (FDR). The Support Vector Machine (SVM) is used to classify the volumes using the feature set so obtained. To handle the problem of limited data, augmentation is used with dropout to prevent over-fitting. The TGA is used to find the number of filters, size of the filters, presence or absence of pooling, and the activation function. This work also explores the applicability of the recently proposed SIREN activations in CNN.[27] In one of our earlier works, a manually crafted CNN was created in which the number of filters, and the number of units in the fully connected layers were determined using empirical analysis. The maximum accuracy obtained from that network was 0.8897, with sigmoid activation. The activation functions ReLU, tanh, and sin yielded an accuracy of 0.9163, 0.8532, and 0.9210, respectively, thus suggesting the role of activation functions on the performance of a model.

In the first experiment, GA was used to find the number of filters in the two convolutional layers, their size, and whether pooling is required. In addition, the number of units in the two fully connected layers was also decided using GA. An accuracy of 0.95017 was obtained in 16 generations. This was followed by the application of the DGA for finding these hyper-parameters. The application of Ng-Wong[31] scheme resulted in an accuracy of 0.95248. The application of Ryan[32] and Yang schemes of DGA[46] yielded accuracies of 0.9569 and 0.96296, respectively, in fewer generations (the last one in 11 generations). The proposed TGA gave an accuracy of 0.9796 in seven generations. The above experiments were conducted for mutation rate 0.02 and crossover rate 0.40. The comparison of the accuracy of the proposed model with the existing models is presented in Table 2. It can be observed from the table that the proposed method performs well in terms of accuracy for the MCI-C versus MCI-NC data compared to the state-of-the-art.

From the above results the following points can be concluded:

- TGA performs better than DGA.
- Among different variants of DGA, the Yang scheme results in the best performance.
- DGA, in general, performs better compared to GA, for the classification of MCI.
- Choice of activation functions affects the performance of a model.
- For this data, the sin activation gives the best performance.
- The choice of hyper-parameters affects the recital of a CNN architecture.

However, this study is carried out at a particular crossover rate and mutation rate. Further, the coarse structure of the CNN was chosen after empirical analysis. The hierarchical TGA-CNN being developed will find the coarse structure and the parameters of the GA as well. The above method is also applied for the classification of MCI and AD patients. The results suggest that the method can be used for the classification of MCI, in general.

3 | DETAILED METHODS

GAs are heuristic search algorithms based on the theory of natural selection.[39] The simple GA uses the brilliance of biological genetics and depends on crossover, mutation, and reproduction with a slender disparity in the algorithm meta-structure. In GA, the fitter chromosomes have greater representation in future generations. They have been extensively used for optimization and selection problems and have demonstrated their superiority time and again. However, these algorithms have been proved better for solving a problem in which the fitness evaluation is measurable against a schedule of static costs but problems with dynamic fitness evaluations have been only diffidently advanced with GA.[40,41]

In nature, the cells contain different sets of chromosomes, which is called ploidy.[42] In the case of a pair of chromosomes, a characteristic of the cells given by an allele is determined based on the phenotype using the dominant or recessive character of the respective gene, and the genetic information thus determined will be transmitted to the offspring.[42] This leads to the transfer of greater diversity of characteristics to the offspring in complex organisms. The prevalent versions of GAs, however, use haploid representation, which retains only one set of each gene, possibly because of ease of implementation.

Goldberg observed that diploid populations were able to adapt to changing environments more readily than haploid populations.[41] For diploid-to-haploid conversion, a dominance scheme is used, one of which is by the triallelic map created by Hollstein.[43] In the scheme, the cardinality of the population is three and the dominance scheme generates a 1 if the sum of alleles is 2 or greater; otherwise, a 0 is generated (Table 3).

TABLE 2 Comparison of performance of the proposed method with existing works

Machine learning methods		Deep learning methods	
Method	**Accuracy (%)**	**Method**	**Accuracy (%)**
Gerardin et al. (2009)[9]	83.00	Basaia et al. (2019)[33]	75.10
Chupin et al. (2009)[10]	64.21	Suk et al. (2017)[34]	74.82
Carlton Chu et al. (2011)[11]	65.00	Li et al. (2015)[26]	57.40
Dai et al. (2013)[12]	71.04	Singh et al. (2017)[35]	72.47
Chong-Yaw Wee et al. (2013)[13]	71.67	Lu et al. (2018)[36]	82.93
Tong et al. (2014)[14]	72.00	Lu et al. (2018)[37]	75.44
Ahmed et al. (2015)[15]	68.72	Ortiz et al. (2016)[38]	82.00
Liu et al. (2018)[16]	72.08	Proposed model	97.961

TABLE 3 Hollstein's triallelic map

_	0	1	2
0	0	0	1
1	0	1	1
2	1	1	1

TABLE 4 Ng, Wong scheme

_	O	i	0	1
o	0	0/1	0	1
i	0/1	1	0	1
0	0	0	0	0/1
1	1	1	0/1	1

TABLE 5 Ryan scheme

_	A	B	C	D
A	0	0	0	1
B	0	0	0	1
C	0	0	1	1
D	1	1	1	1

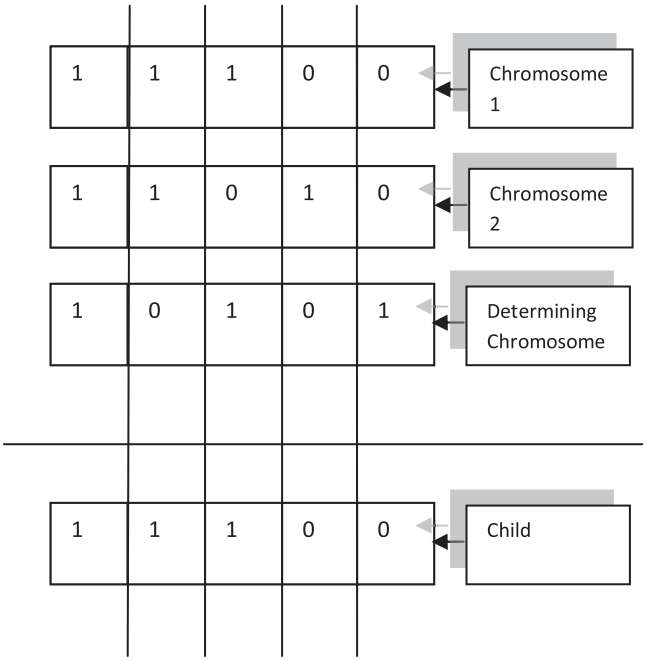

FIGURE 1 The formation of a child chromosome in the proposed dominance scheme

Eshelman and Schaffer were able to establish, using the atmosphere and metabolism problem, the ability of GA with diploidy and dominance to create dynamically stable systems without a loss of fitness.[44] Ng and Wong used four genotypic alleles and employed randomness when both alleles were dominant or both were recessive.[31] That is, dominant alleles "0" and "1" may produce a "0" or a "1" (Table 4). The major problem in this approach was uncertainty. Ryan proposed an additive dominance scheme in which genotypic alleles are represented by ordered values that are combined using pseudo-arithmetic to determine the phenotypic allele.[32] The dominance map proposed by Ryan is shown in Table 5. Yang and Yao[45] proposed an adaptive dominance mechanism. According to them the cardinality of the genotypic alleles and the uncertainty in the dominance scheme are the two key factors that affect the performance of the technique.[45] They carried out experiments with the help of a tool developed in one of Yang's earlier works[46,47] and proved that dominance scheme is better than previous ones.

In addition to haploids and diploids, triploids also exist. For example, a banana has three sets of chromosomes, two sets from one parent and one set from the other parent. Triploids seldom produce eggs or sperm that have a balanced set of chromosomes and so a successful seed set is very rare.[48] Note that bananas are parthenocarpic, that is, fruit developed without fertilization. The parthenocarpic fruits have some remarkable properties, such as longer shelf life.

The proposed technique of obtaining a chromosome from three chromosomes is depicted in Figure 1. The process and the corresponding formula have been stated as follows. The proposed dominance technique uses three chromosomes to determine the child. In the discussion

TABLE 6 The proposed dominance scheme

Chromosome 1	Chromosome 2	Determining chromosome	Child
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	1

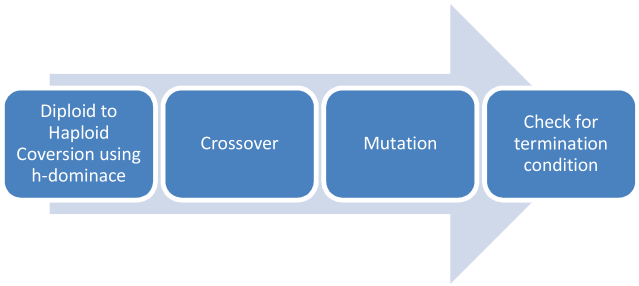

FIGURE 2 The proposed dominance scheme based diploid genetic algorithm

that follows, the first parent is referred to as "chromosome 1," the second as "chromosome 2." The chromosome that determines the dominance is referred to as "the determining chromosome." If the cells in both the chromosomes are "1," the corresponding child would be "1," irrespective of the value of the cell in the determining chromosome. Likewise, if both the cells are "0," the cell in the child would be "0." In case the two cells are "1"and "0," the output would be determined by the "determining chromosome." Value "1" in the determining chromosome would result in a "1," whereas "0" in the determining chromosome would lead to a "0" in the child. The possible combinations are shown in Table 6.

The bit in the child chromosome (C) can, therefore, be expressed in terms of the corresponding bit in chromosome 1 (C1), chromosome 2 (C2), and determining chromosome (D) as follows (Formula 1).

$$C = C_1.C_2 + C_1.D + C_2.D \quad (1)$$

The above schema is used to convert triploid to haploid in TGA. The steps of TGA are as follows. First, some random chromosomes are generated. For example, if initially 3n chromosomes are taken, then n chromosomes are obtained after applying the dominance technique. The process of crossover is then applied. This is followed by mutation. Finally, the "fit" individuals proceed to the next generation, after selection. The process is depicted in Figure 2.

The proposed scheme does not yield uncertain results and therefore handles the problem pointed out by Yang; also, the cardinality is low compared to the others.

Given the coarse structure of a CNN (the number of convolutional layers and the number of dense layers), the work aims to find the

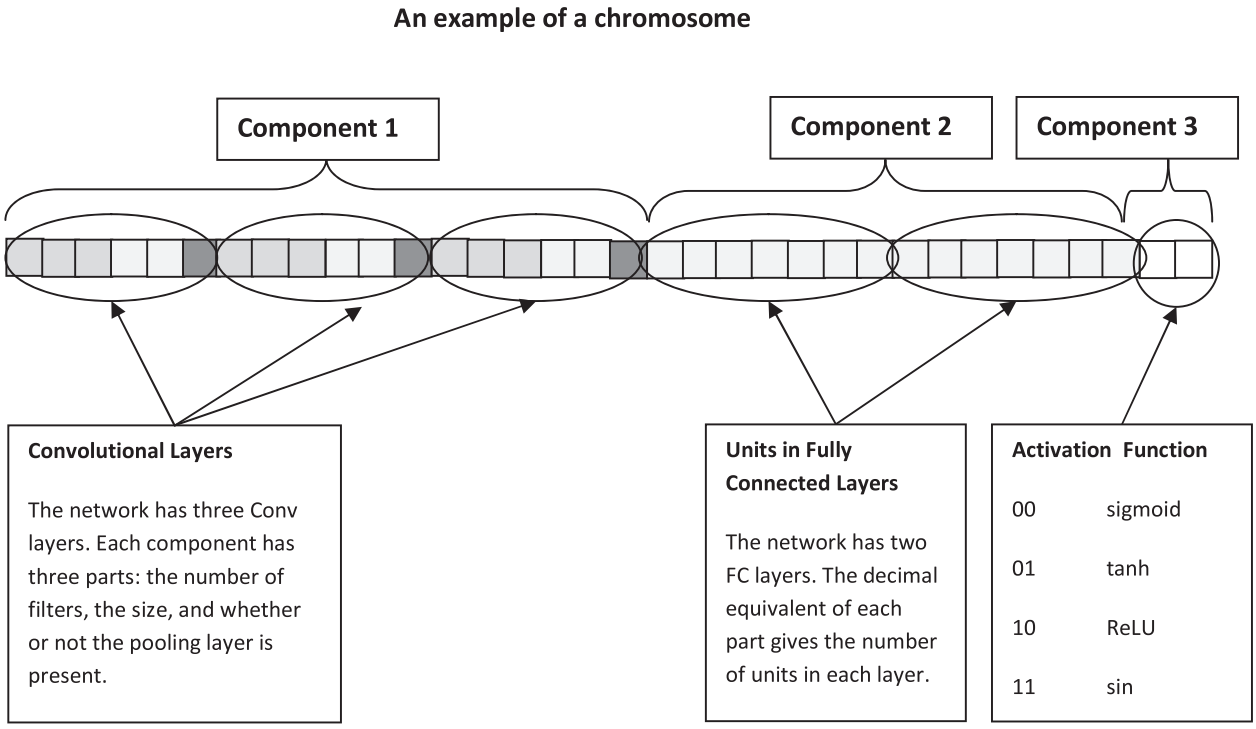

FIGURE 3 An example of a chromosome used in this work. FC, fully connected

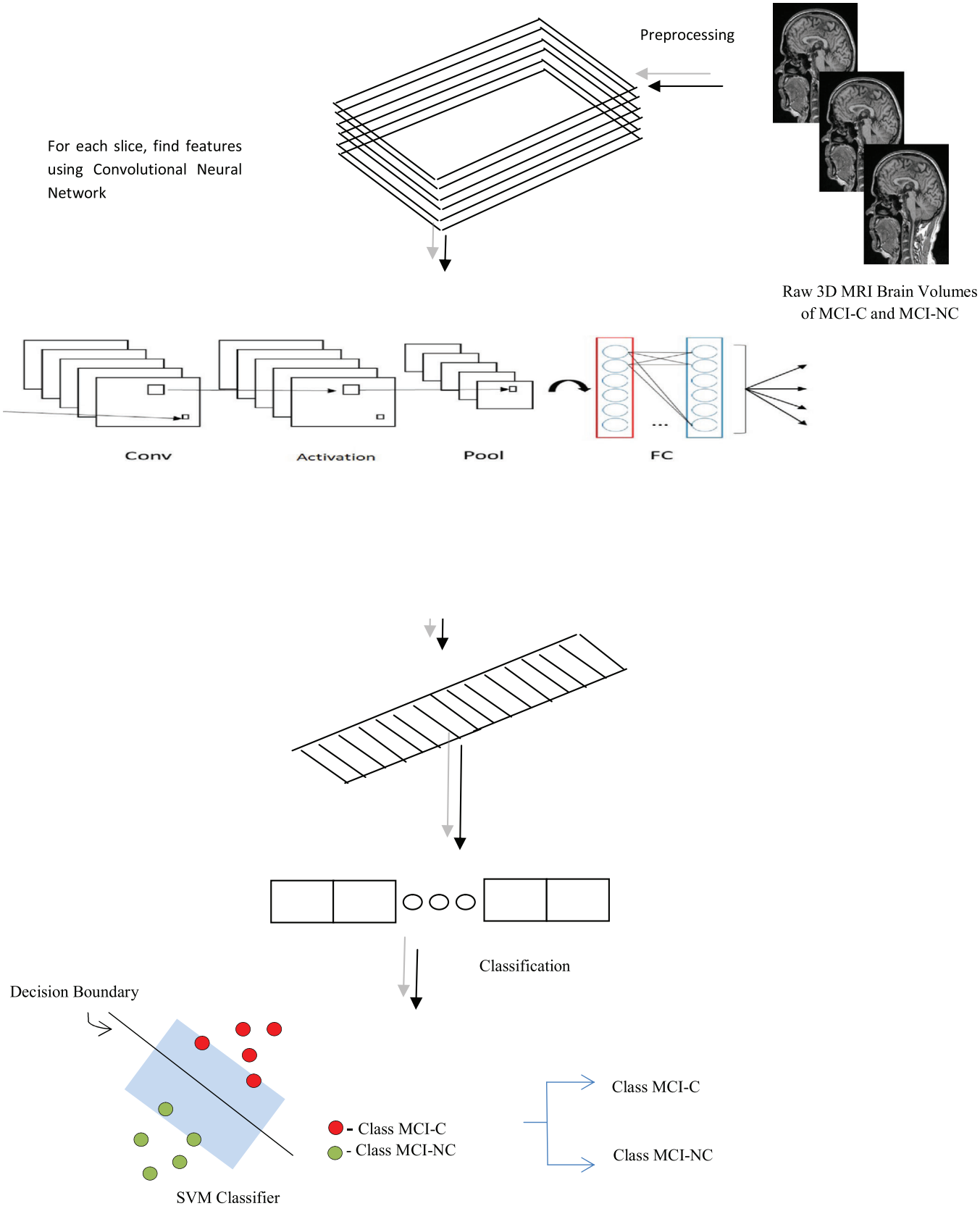

FIGURE 4 Steps for finding accuracy for each chromosome. MCI-C, mild cognitive impairment converts; MCI-NC, mild cognitive impairment non-converts; MRI, magnetic resonance imaging

number of filters, their size, the presence or absence of pooling layers, and the activation functions using GA and its variants. This work uses chromosomes represented as a binary string. A chromosome has three components: the first component gives information regarding the convolutional layers, the second component the number of units in the fully connected layers, and the third discloses the activation function. Further, the first component has n_c parts where n_c is the number of convolutional layers. Each part has three subparts: the first tells the number of filters, the second their size, and the third represents the presence or absence of the pooling. The second component has n_{fc} parts, where n_{fc} is the number of fully connected layers. Each part tells the number of units in the fully connected layers. The last component represents the activation function.

It may be noted that for each part the maximum and minimum limit would determine the number of cells. For example, if $n_1 \leq n_{\text{filters in convlayer}} \leq n_2$, then find the smallest n so that, $n_2 - n_1 \leq 2^n$. The number of cells in the subpart will be n.

The decoding of this cell would be done as follows:

$$out = \left\lfloor n_1 + \frac{\text{the decimal equivalent of the value in the subpart}}{2^n - 1} \times (n_2 - n_1) \right\rfloor$$

where, $\lfloor\ \rfloor$ is the greatest integer function. Figure 3 depicts an example of a chromosome used in the work. In the example, the number of convolutional layers is three and the number of dense layers is two.

This work uses the roulette wheel selection. The experiments are conducted using 1-point and 2-point crossover. The purpose of mutation is to prevent the GA from converging to local optima, but if the mutation rate is high, the GA is changed to random search.[39]

The fitness function used in the work is as follows:

$$fitness = k \times accuracy.$$

The proposed framework is summed up as follows.

For the given data:

1. Divide the data into train and test sets.
2. Generate the initial population of TGA and craft the CNN architecture accordingly.
3. For each generation, carry out the following steps for each MRI volume in the train set:
   a. For each slice of the MRI volume, obtain the feature vector using the output of the penultimate layer of the above network.
   b. Concatenate the feature vectors of each slice horizontally to obtain the feature vector of the volume. This would be used for the classification using the SVM.
4. Place the FDR) values of the features obtained from step 3 in the decreasing order. The testing phase would use the indices so obtained.
5. Use the SVM to train the model using the train set.
6. The procedure stated in step 3 is then applied for the test-set. Use the indices obtained in step 4 to obtain the relevant features.
7. Obtain the accuracy of the proposed method.
8. Apply crossover and mutation operations and for the new population carry out steps 3 to 7, until the termination condition is reached.

The steps of finding fitness for each chromosome, are shown in Figure 4. The results have been summarized in Table 7.

TABLE 7 Performance of heuristic search algorithms in MCI-C ad MCI-NC classification

Algorithm	Accuracy	Generations
Triploid genetic algorithms	0.97961	7
Yang scheme	0.96296	11
Ryan scheme	0.95691	14
Ng-Wong	0.95248	14
Simple genetic algorithm	0.95017	16

Abbreviations: MCI-C, mild cognitive impairment converts; MCI-NC, mild cognitive impairment non-converts.

4 | CONCLUSION

The performance of a CNN depends on the choice of hyper-parameters. Manually choosing these may not result in the optimal architecture. This work uses a proposed variant of GA called TGA to find the parameters of a CNN. The role of the activation function, particularly SIREN activation, on the performance of the network is also explored in the work. The performance of the proposed model is better compared to the existing models. However, the results depend on the choice of crossover rate and mutation rate. Also, if the volume of the input tensor is reduced, while retaining most of the information, the computation time of the proposed algorithm will be drastically reduced. Moreover, the depth of the network may also affect the results but may lead to vanishing gradients. Finally, the concept of ploidy can be further explored.

For the first part, we aim to develop a hierarchical GA that finds the parameters and the hyper-parameters of the given CNN for the classification of a given dataset. To reduce the volume of the input tensor, we have developed a principal component analysis–CNN–based model that reduces the input volume while retaining the information regarding the correlation between the slices of the brain. The problem of vanishing gradient in the model will be handled using a parallel model like in the case of inception. The future work aims to develop a model that reduces the volume of the input tensor, while retaining most of the information, and to use a parallel-inception type network, to classify MCI. Hierarchical TGA will be used to choose the parameters of the network.

ACKNOWLEDGMENTS

Data used for this study were downloaded from the ADNI Repository (http://loni.adni.com). Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI;

National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie; Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC; Johnson & Johnson Pharmaceutical Research & Development LLC; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. The first author has a Visvesvaraya Fellowship from the Ministry of Electronics and Information Technology, Government of India. No funds were received for the publication of this article and hence waiver was applied.

CONFLICTS OF INTEREST

The authors have no conflicts of interest.

ORCID

Harsh Bhasin https://orcid.org/0000-0001-8151-0501

REFERENCES

1. https://alzcnfl.wordpress.com/2016/04/28/alzheimers-a-real-love-story/ Accessed April 25, 2021.
2. The Dementias: Hope Through Research | National Institute of Neurological Disorders and Stroke. https://www.ninds.nih.gov/Disorders/Patient-Caregiver-Education/Hope-Through-Research/Dementia-Hope-Through-Research. Accessed January 1, 2019.
3. World Alzheimer Report 2018, The state of the art of dementia research: New (2018). Alzheimer's disease international, Chicago. https://www.alzint.org/u/WorldAlzheimerReport2018.pdf
4. Mild Cognitive Impairment, https://alz.org/alzheimers-dementia/what-is-dementia/related_conditions/mild-cognitive-impairment, Accessed January 1, 2019.
5. Geda YE. Mild cognitive impairment in older adults. *Current Psychiatry Reports*. 2012;14(4):320-327.
6. Ward A, Tardiff S, Dye C, Arrighi H. Rate of conversion from prodromal Alzheimer's disease to Alzheimer's dementia: a Systematic Review of the Literature. *Dement Geriatr Cogn Disord*. 2013;3(1):320-332.
7. Suk H-I, Shen D, Deep learning-based feature representation for AD/MCI classification. *The International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2013; 583-590.
8. Singh S, et al. Deep Learning based Classification of FDG-PET Data for Alzheimers Disease Categories. *Proc. SPIE– Int. Soc. Opt. Eng.*; 2017: 10572.
9. Gerardin E, Chételat G, Chupin M, et al. Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging. *Neuroimage*. 2009;47(4):1476-1486. 2009.
10. Chupin M, Gérardin E, Cuingnet R, et al. Fully automatic hippocampus segmentation and classification in Alzheimer's disease and mild cognitive impairment applied on data from ADNI. *Hippocampus*. 2009;19(6):579-587.
11. Chu C, Hsu Ai-L, Chou K-H, Bandettini P, Lin C. Does feature selection improve classification accuracy? Impact of sample size and feature selection on classification using anatomical magnetic resonance images. *Neuroimage*. 2012;60(1):59-70.
12. Dai D, He H, Vogelstein JT, Hou Z. Accurate prediction of AD patients using cortical thickness networks. *Mach Vis Appl*. 2013;24(7):1445-1457.
13. Wee C-Y, Yap P-T, Shen D. Prediction of Alzheimer's disease and mild cognitive impairment using baseline cortical morphological abnormality patterns. *Hum Brain Mapp*. 2013;34(12):3411-3425.
14. Tong T, Wolz R, Gao Q, Guerrero R, Hajnal JV, Rueckert D. Multiple instance learning for classification of dementia in brain MRI. *Med Image Anal*. 2014;18(5):808-818.
15. Ben Ahmed O, Benois-Pineau J, Allard M, Ben Amar C, Catheline G. Classification of Alzheimer's disease subjects from MRI using hippocampal visual features. *Multimed Tools Appl*. 2015;74(4):1249-1266.
16. Liu J, Li M, Lan W, Wu F-X, Pan Yi, Wang J. Classification of Alzheimer's disease using whole brain hierarchical network. *IEEE/ACM Trans Comput Biol Bioinform*. 2018;15(2):624-632.
17. Bhasin H, Agrawal RK. A combination of 3-D discrete wavelet transform and 3-D local binary pattern for classification of mild cognitive impairment. *BMC Med Inform DecisMak*. 2020;20:1-10.
18. Zhang H, Sachdev PS, Wen W, et al. Gray matter atrophy patterns of mild cognitive impairment subtypes. *J Neurol Sci*. 2015;315(1-2): 26-32.
19. Texture Analysis Using the Gray-Level Co-Occurrence Matrix (GLCM) - MATLAB & Simulink - MathWorks United Kingdom.
20. Ojala T, Pietikainen M, Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans Pattern Anal Mach Intell*. 2002;24(7):971-987.
21. Ojala T, Pietikäinen M, Harwood D. A comparative study of texture measures with classification bafsed on featured distributions. *Pattern Recognit*. 1996;29(1):51-59.
22. Ojala T, Pietikäinen M, Mäenpää T, A generalized local binary pattern operator for multiresolution gray scale and rotation invariant texture classification, *The International Conference on Advances in Pattern Recognition*, 2001; 399-408.
23. Mallat SG. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans Pattern Anal Mach Intell*. 1989;11(7):674-693.
24. Lecun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521:436-444.
25. Convolutional Neural Networks for Visual Recognition, Available: https://cs231n.github.io/convolutional-networks/ [Accessed: 25-Feb-2021].
26. Li F, Tran L, Thung KH, Ji S, Shen D, Li J. A robust deep model for improved classification of AD/MCI patients. *IEEE J Biomed Health Inform*. 2015;19:1610-1616.
27. Sitzmannet V. Implicit Neural Representations with Periodic Activation Functions. ArXiv abs/2006.09661, 2020. https://arxiv.org/pdf/2006.09661.pdf
28. Petersen RC, Aisen PS, Beckett LA, et al. Alzheimer's Disease Neuroimaging Initiative (ADNI): clinical characterization. *Neurology*. 2010;74(3):201-209.

29. Salvatore C, Cerasa A, Battista P, Gilardi MC, Quattrone A, Castiglioni I. Magnetic resonance imaging biomarkers for the early diagnosis of Alzheimer's disease: a machine learning approach. *Front Neurosci*. 2015;9.
30. Friston KJ, Holmes AP, Worsley KJ, Poline J-P, Frith CD, Frackowiak RSJ. Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapp*. 1994;2(4):189-210. https://onlinelibrary.wiley.com/doi/epdf/10.1002/hbm.460020402
31. Ng KP, Wong KC, A New Diploid Scheme and Dominance Change Mechanism for Non-Stationary Function Optimization, *In the Proceedings of the 6th International Conference on Genetic Algorithms*, Eshelman L J (Ed.), Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1994, pp. 159-166.
32. Ryan C, Degree of Oneness, in the proceedings of the 1994 *ECAI workshop of Genetic Algorithm*, 1994.
33. Basaia S, Agosta F, Wagner L, et al. Alzheimer's Disease Neuroimaging Initiative, Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. *Neuroimage Clin*. 2018;21:101645. Epub.
34. Suk H-II, Lee S-W, Shen D. Deep ensemble learning of sparse regression models for brain disease diagnosis. *Medical Image Anal*. 2017;37:101-113.
35. Singh S, Srivastava A, Mi L, et al., Deep-learning-based classification of FDG-PET data for Alzheimer's disease categories, in *Proceedings of the 13th International Conference on Medical Information Processing and Analysis*, (Bellingham, WA: International Society for Optics and Photonics), 2017. https://doi.org/10.1117/12. 2294537.
36. Lu D, Popuri K, Ding GW, Balachandar R, Beg MF. Multimodal and multiscale deep neural networks for the early diagnosis of Alzheimer's disease using structural MR and FDG-PET images. *Sci Rep*. 2018(a);8:5697.
37. Lu D, Popuri K, Ding GW, Balachandar R, Beg MF. and Alzheimer's Disease Neuroimaging Initiative, Multiscale deep neural network based analysis of FDG-PET images for the early diagnosis of Alzheimer's disease. *Med Image Anal*. 2018(b);46:26-34. https://doi.org/10.1016/j.media.2018.02.002.
38. Ortiz A, Munilla J, Gorriz JM, Ramirez J. Ensembles of deep learning architectures for the early diagnosis of the Alzheimer's disease. *Int J Neural Syst*. 2016;26:1650025.
39. Holland JH. *adaptation in Natural and Artificial Systems*. University of Michigan Press; 1975.
40. Kenneth A, De J. *Evolutionary Computing, A Unified Approach*. MIT Press; 2006. https://mitpress.mit.edu/books/evolutionary-computation
41. Goldberg DE. *Genetic Algorithms in Search, Optimization and Machine Learning*. 1st ed. Addison-Wesley Longman Publishing Co., Inc.; 1989.
42. Hartl D, (2011). Essential Genetics: A Genomics Perspective. Jones & Bartlett Learning. p. 177. ISBN 978-0-7637-7364-9
43. Hollstein RB. *Artificial Genetic Adaptation in Computer Control Systems*. University of Michigan; 1971. https://www.proquest.com/openview/1681d71f520d7d39eb12156d59b763ed/1?pqorigsite=gscholar&cbl=18750&diss=y
44. Eshelman LJ, Schaffer JD. Real-coded genetic algorithms and interval-schemata. *Found Genetic Alg*. 1993;2:187-202.
45. Yang S, Yao X. Experimental study on population based incremental learning algorithm for dynamic optimization problem. *Soft Computing*. 2008:815-834.
46. Yang S. On the design of diploid genetic algorithms for problem optimization in dynamic environments. In Proceedings of IEEE International Conference on Evolutionary Computation, 2006; 1362-1369.
47. Yang S, Yao X. Experimental study on population-based incremental learning algorithms for dynamic optimization problems. *Soft Computing*. 2005;9:815-834.
48. Burr B, Burr F. How do seedless fruits arise and how are they propagated? *Sci Am*. 2020.
49. Sarraf S & Tofighi G for the Alzheimer's Disease Neuroimaging Initiative, DeepAD: Alzheimer's Disease Classification via Deep Convolutional Neural Networks using MRI and fMRI, 2016; 1-15.

How to cite this article: Bhasin H, Agrawal RK. Triploid genetic algorithm for convolutional neural network–based diagnosis of mild cognitive impairment. *Alzheimer's Dement*. 2021;1-9. https://doi.org/10.1002/alz.12565