Electronics Letters

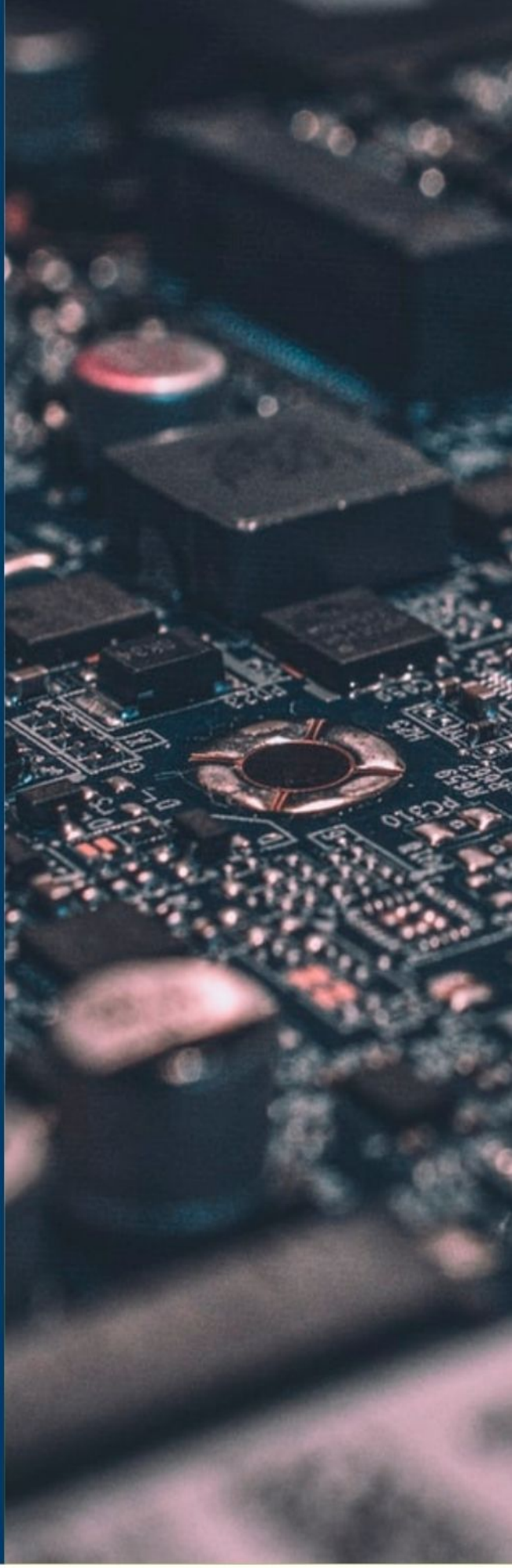
Special issue Call for Papers

Be Seen. Be Cited.
Submit your work to a new IET special issue

Connect with researchers and experts in your field and share knowledge.

Be part of the latest research trends, faster.

Read more



The Institution of Engineering and Technology

Check for updates

Brain MRI-based Wilson disease tissue classification: an optimised deep transfer learning approach

L. Saba, M. Agarwal, S.S. Sanagala, S.K. Gupta, G.R. Sinha, A.M. Johri, N.N. Khanna, S. Mavrogeni, J.R. Laird, G. Pareek, M. Miner, P.P. Sfikakis, A. Protogerou, V. Viswanathan, G.D. Kitas and J.S. Suri✉

Wilson's disease (WD) is caused by the excessive accumulation of copper in the brain and liver, leading to death if not diagnosed early. WD shows its prevalence as white matter hyperintensity (WMH) in MRI scans. It is challenging and tedious to classify WD against controls when comparing visually, primarily due to subtle differences in WMH. This Letter presents a computer-aided design-based automated classification strategy that uses optimised transfer learning (TL) utilising two novel paradigms known as (i) MobileNet and (ii) the Visual Geometric Group-19 (VGG-19). Further, the authors benchmark TL systems against a machine learning (ML) paradigm. Using four-fold augmentation, VGG-19 is superior to MobileNet demonstrating accuracy and area under the curve (AUC) pairs as *95.46 ± 7.70%, 0.932* ($p < 0.0001$) and *86.87 ± 2.23%, 0.871* ($p < 0.0001$), respectively. Further, MobileNet and VGG-19 showed an improvement of *3.4* and *13.5%*, respectively, when benchmarked against the ML-based soft classifier – Random Forest.

Introduction: Wilson's disease (WD) is a rare disease, which causes copper accumulation in the liver and brain. As per National Organizations for Rare Disorders (NORD), WD affects 1 in every 30,000–40,000 people in the world [1]. If not diagnosed early, WD can lead to disability and death [2]. The disease can further lead to neurological disorders such as Alzheimer's and Parkinson's diseases [3].

WD can be identified and diagnosed using MRI as it reflects the brain to have white matter hyperintensity (WMH) [4–10]. However, due to subtle changes in the soft tissues of brain white matter and grey matter convolutions, it is challenging and tedious to classify visually.

WD classification problems can be looked into by a very popular paradigm such as artificial intelligence (AI). In this field of computer science, machine learning (ML) has dominated, especially in healthcare applications which includes prostate [11, 12], ovarian [13], liver [14, 15], thyroid [16–18], skin [19], diabetes [20], gene [21], heart [22], coronary [23], signal [24, 25] and carotid atherosclerotic plaque [26–32]. These methods have produced accuracies ranging in the high 80s or even 90s, but the painstaking effort has been put for feature extraction and selection, thereby manually optimising them, leading to an *ad hoc* solution.

Deep learning (DL), on the contrary, is a 'class' of AI that has recently revolutionised image classification framework [33, 34]. Even more recently, the paradigm of transfer learning (TL) brings a reduction in training time which is deep rooted in conventional DL frameworks [35]. Since the crux of the classification process seems to be partially resolved both in terms of feature extraction and prevention of training time, we can hypothesise that our WD computer-aided design system consisting of (a) MobileNet and (b) Visual Geometric Group-19 (VGG-19) will be more accurate compared to existing ML systems.

Methodology: Our cohort consist of 46 (37 controls, 9 diseased) patients T2W-TSE MRI scans with a mean age of 40.73 + 11.3 years with equal male and female ratio. All the scans were obtained from Azienda Ospedaliero Universitaria (A.O.U.) by taking the approvals from the institutional ethics committee and patient's approval. The cohort consisted of full scans of MRI WMH images, so it is essential to segment the images and remove the skull part for classification. Using 'dcm2niix' software, we converted all the DICOM images to niix for loading to 'brain suite' [36]. This helped in removing the skull part. Fig. 1 shows the control and WD sample scans of six patients and their corresponding segmented (skull stripping) images.

Augmentation: Since our cohort consisted of 37 controls and nine WD patients, and each patient had 12–13 MRI slices, totalling to 458 controls and 115 WD slices. Hence, clearly, it was a 'class imbalance' 'problem. As the TL paradigm in the deep framework is well adaptable for large sample sizes, we, therefore, used the 'augmentor' API to generate a balanced cohort with 458 images in each class. Further, we augmented the data two times (916 images per class) for superior performance (Table 1).

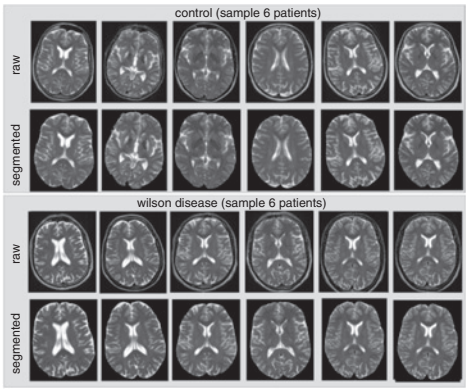


Fig. 1 *Sample images of controlled and diseased MRI scans with corresponding segmented images (skull stripping)*

Table 1: Details of the augmented cohort

SN	Augmentation folds	No. of images per class
1	balanced	458 (equal control and WD)
2	augmented 2× (Augm 2×)	916 (equal control and WD)
3	augmented 3× (Augm 3×)	1378 (equal control and WD)
4	augmented 4× (Augm 4×)	1832 (equal control and WD)

Transfer learning architectures and experimental protocol: The proposed (i) MobileNet and (ii) VGG-19 are shown in Fig. 2. The MobileNet consisted of six pools of convolutional blocks with an average pooling layer followed by one fully connected layer. The VGG-19 consisted of 19 layers, 4 pools of convolution blocks and 5 max-pooling layers followed by 3 fully connected layers with the output softmax layer.

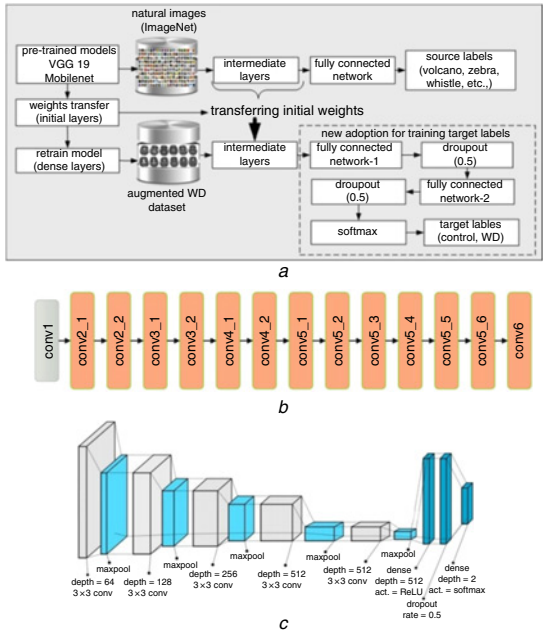


Fig. 2 *Architecture of*

a Proposed TL
b MobileNet
c VGG-19

TL is applied to pre-trained neural networks, in which weights already learned on different data sets are further calibrated for a target data set. In this study, the pre-trained networks VGG19 and MobileNet were first trained on an 'ImageNet' data set having 1000 classes consisting of images such as a zebra, volcano, whistle etc. Since the intended classification has only two classes hence final Softmax layer is updated at the end of the entire network and retrained for a few epochs to reshape the weights for the WD data set. To test the updated convolution neural network, the K10 protocol is used (90% training and 10% testing). We benchmarked our proposed TL models with ML models *K*-nearest

neighbour, Decision Tree and Random Forest. Haralick features were extracted from the segmented MRI scans.

Results: We trained and tested our TL models on augm1×, 2×, 3×, 4× data sets using K10 cross-validation achieving a classification accuracy and receiver operating characteristics (ROC) for MobileNet and VGG-19 as 86.87%, 0.868 ($p < 0.0001$) and 95.46%, 0.954 ($p < 0.0001$), respectively. We benchmarked our proposed TL methods with previously developed ML techniques.

Further, we benchmarked against three ML models with the best accuracy/AUC combination for RF classifier as 83.04 + 2.81%, 0.834 ($p < 0.0001$) (Fig. 3).

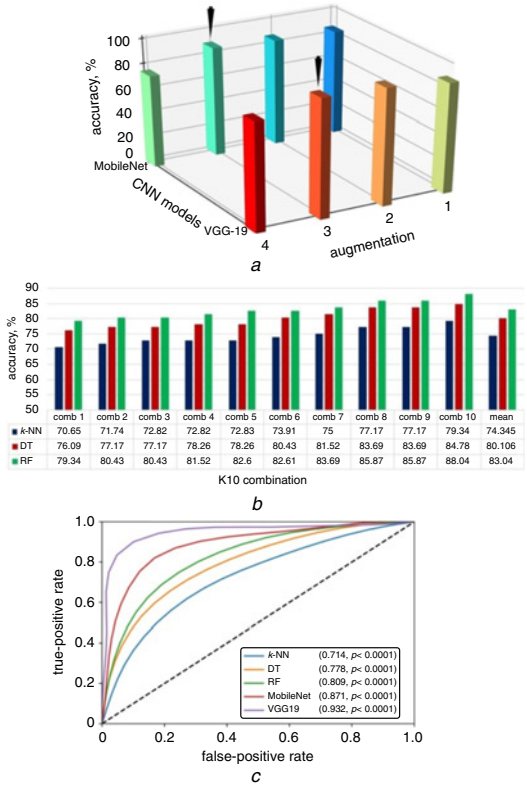


Fig. 3 *Visualisation of the performance analysis*

a Three-dimensional representation of optimisation of TL methods (augmentation versus accuracy for MobileNet and VGG-19)
b ML classifiers with K10 cross-validation protocol [accuracy of ML classifiers (*K*-NN versus DT versus RF)]
c ROC curves of proposed five AI methods

Sample size computation: Sample size determination was accomplished using (1), based on the mean difference of TL models while keeping the augmentation 2× constant. With type I error of 1% $Z_\alpha = 3.2905$, $Z_{1-\beta} = 1.6449$ for type II error with 1%

$$\text{Sample size} = \frac{2 \times \left(Z_\alpha + Z_{1-\beta}\right)^2 \times \sigma^2}{\Delta^2} \quad (1)$$

One would require a cohort of size 793. The cohort adopted in this study was 1832. Thus, we well met the requirements of the cohort size by 56.71%.

Performance evaluation and scientific validation: Our optimised TL system showed that it had an accuracy compared with the ML-based system by *13.01%*. Thus, with the help of the TL model, we reduced one of the major limitations of DL, which is due to computational cost. We evaluated our model on the most widely and well-accepted data sets consisting of facial biometric [37] and animal data (ASSIRA) [38] with optimal performance of classification accuracy of 98.27 and 93.45%, respectively.

Discussion: The objective of this Letter was to optimise the TL models for detecting WD by using MobileNet and VGG-19 approaches. We benchmarked our TL models against other studies as shown in Table 2. This is the first study of its kind to use TL for classification of WD. This Letter did not consider the tissue characterisation component in the analysis and our focus was only on TL classification and its' benchmarking against ML-based models.

Table 2: Benchmarking the proposed method against previous studies

SN	C1	C3	C5	C6
	Author, Reference (year)	Application, technique, modality	AI	ACC (%) (AUC)
R1	Kaden *et al.* [39] (2015)	WD, SVM, and PGLVQ	ML	90.1
R2	Suk *et al.* [40] (2017)	ALZ[a], DBN + NN, MRI (ADNI)	DL	85.91 (0.91)
R3	Abiwinanda *et al.* [41] (2018)	BT[c], CNN, MRI	DL	84.19
R4	Zhang *et al.* [42] (2019)	MCI[b], SSGSR fMRI	ML	88.50 (0.965)
R5	Abrol *et al.* [43] (2018)	MCI[b] versus ALZ[a] ResNet, MRI	DL	82.7 (0.89)
R6	Jing *et al.* [44] (2019)	WD, SVM and ICA, fMRI	ML	89.4 (0.94)
R7	Richhariya *et al.* [45] (2020)	MCI[b], SVM, MRI	ML	90
R9	Liu *et al.* [46] (2020)	MCI[b], SVM, MRI	ML	88.5 (0.90)
R10	MobileNet	CNN	DL	86.87 (0.871)
R11	VGG-19	CNN	DL	95.46 (0.932)

[a] ALZ: Alzheimer's Disease, [b] MCI: mild cognitive impairment, [c] BT: Brain Tumour.

Limitations: The main strength of the TL system was the saving of computational time compared with ML approaches. Furthermore, we showed the implementation of five kinds of AI models (two kinds of TL and three kinds of ML models). The main limitation of this study was the small sample size requiring augmentation procedures. A larger data set would better support the learning of more complex patterns for enhanced diagnostic performance of the algorithm. This pilot study could be improved by: (i) usage of advanced texture feature extraction systems [47], (ii) usage of a reinforcement model to avoid unbalanced data, (iii) superior segmentation models such as scholastic models [48] and (iv) benchmarking against other DL strategies [49, 50].

Conclusion: We proposed the optimised TL methods for the classification of WD. We achieved an accuracy of *95.46* and *86.87%* and our model performed better than ML methods by *13.01* and *4.02%*. We validated our model with a power study and the most widely accepted data sets.

© The Institution of Engineering and Technology 2020
Submitted: *16 July 2020* E-first: *4 November 2020*
doi: 10.1049/el.2020.2102
One or more of the Figures in this Letter are available in colour online.

L. Saba (*Department of Radiology, Azienda Ospedaliero Universitaria (A.O.U.), Cagliari, Italy*)

M. Agarwal and S.K. Gupta (*CSE Department, Bennett University, Greater Noida, UP, India*)

S.S. Sanagala (*Department CSE, CMR College of Engineering & Technology, Hyderabad, India*)

G.R. Sinha (*Myanmar Institute of Information Technology (MIIT), Mandalay, Myanmar*)

A.M. Johri (*Department of Medicine, Division of Cardiology, Queen's University, Kingston, Ontario, Canada*)

N.N. Khanna (*Department of Cardiology, Indraprastha APOLLO Hospitals, New Delhi, India*)

S. Mavrogeni (*Cardiology Clinic, Onassis Cardiac Surgery Center, Athens, Greece*)

J.R. Laird (*Heart and Vascular Institute, Adventist Health St. Helena, St Helena, CA, USA*)

G. Pareek (*Minimally Invasive Urology Institute, Brown University, Providence, Rhode Island, USA*)

M. Miner (*Men's Health Center, Miriam Hospital Providence, Rhode Island, USA*)

P.P. Sfikakis (*Rheumatology Unit, National Kapodistrian University of Athens, Athens, Greece*)

A. Protogerou (*Department of Cardiovascular Prevention, National and Kapodistrian University of Athens, Athens, Greece*)

V. Viswanathan (*MV Hospital for Diabetes & Professor M Viswanathan Diabetes Research Centre, Chennai, India*)

1350911x, 2020, 25, Downloaded from https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/el.2020.2102 by CochraneItalia, Wiley Online Library on [26/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

G.D. Kitas (*R & D Academic Affairs, Dudley Group NHS Foundation Trust, Dudley, United Kingdom*)

J.S. Suri (*Stroke Monitoring and Diagnosis Division, AtheroPoint™, Roseville, CA, USA*)

✉ E-mail: Jasjit.Suri@AtheroPoint.com

References

1 Dusek, P, Litwin, T., and Czlonkowska, A.: 'Wilson disease and other neurodegenerations with metal accumulations', *Neurol. Clin.*, 2015, **33**, (1), pp. 175–204
2 Medici, V, Rossaro, L., and Sturniolo, G.: 'Wilson disease—a practical approach to diagnosis, treatment and follow-up', *Dig. Liver Dis.*, 2007, **39**, (7), pp. 601–609
3 Ala, A., Walker, A.P., Ashkan, K., *et al.*: 'Wilson's disease', *Lancet*, 2007, **369**, (9559), pp. 397–408
4 Roberts, E.A.: 'A practice guideline on Wilson disease', *Hepatology*, 2003, **37**, pp. 1475–1492
5 Singh, P., Ahluwalia, A., Saggar, K., *et al.*: 'Wilson's disease: MRI features', *J. Pediatr. Neurosci.*, 2011, **6**, (1), p. 27
6 Parekh, J.R., and Agrawal, P.R.: 'Wilson's disease:'face of giant panda'and 'trident'signs together', *Oxf. Med. Case. Reports.*, 2014, **2014**, (1), pp. 16–17
7 Yousaf, M., Kumar, M., Ramakrishnaiah, R., *et al.*: 'Atypical MRI features involving the brain in Wilson's disease', *Radiol. Case. Rep.*, 2009, **4**, (3), p. 312
8 Saba, L., Lucatelli, P., Anzidei, M., *et al.*: 'Volumetric distribution of the white matter hyper-intensities in subject with mild to severe carotid artery stenosis: does the side play a role?', *J. Stroke Cerebrovasc. Dis.*, 2018, **27**, (8), pp. 2059–2066
9 Saba, L., Sanfilippo, R., Porcu, M., *et al.*: 'Relationship between white matter hyperintensities volume and the circle of Willis configurations in patients with carotid artery pathology', *Eur. J. Radiol.*, 2017, **89**, pp. 111–116
10 Porcu, M., Balestrieri, A., Siotto, P., *et al.*: 'Clinical neuroimaging markers of response to treatment in mood disorders', *Neurosci. Lett.*, 2018, **669**, pp. 43–54
11 Pareek, G., Acharya, U.R., Sree, S.V., *et al.*: 'Prostate tissue characterization/classification in 144 patient population using wavelet and higher order spectra features from transrectal ultrasound images', *Technol. Cancer Res. Treat.*, 2013, **12**, (6), pp. 545–557
12 McClure, P., Elnakib, A., El-Ghar, M.A., *et al.*: 'In-vitro and in-vivo diagnostic techniques for prostate cancer: a review', *J. Biomed. Nanotechnol.*, 2014, **10**, (10), pp. 2747–2777
13 Acharya, U.R., Sree, S.V., Kulshreshtha, S., *et al.*: 'Gynescan: an improved online paradigm for screening of ovarian cancer via tissue characterization', *Technol. Cancer Res. Treat.*, 2014, **13**, (6), pp. 529–539
14 Acharya, U.R., Sree, S.V., Ribeiro, R., *et al.*: 'Data mining framework for fatty liver disease classification in ultrasound: a hybrid feature extraction paradigm', *Med. Phys.*, 2012, **39**, (7Part1), pp. 4255–4264
15 Saba, L., Dey, N., Ashour, A.S., *et al.*: 'Automated stratification of liver disease in ultrasound: an online accurate feature classification paradigm', *Comput. Methods Programs Biomed.*, 2016, **130**, pp. 118–134
16 Acharya, U.R., Swapna, G., Sree, S.V., *et al.*: 'A review on ultrasound-based thyroid cancer tissue characterization and automated classification', *Technol. Cancer Res. Treat.*, 2014, **13**, (4), pp. 289–301
17 Acharya, U., Vinitha Sree, S., Mookiah, M., *et al.*: 'Diagnosis of Hashimoto's thyroiditis in ultrasound using tissue characterization and pixel classification', *Proc. Inst. Mech. Eng., H: J. Eng. Med.*, 2013, **227**, (7), pp. 788–798
18 Acharya, U.R., Sree, S.V., Krishnan, M.M.R., *et al.*: 'Non-invasive automated 3D thyroid lesion classification in ultrasound: a class of ThyroScan™ systems', *Ultrasonics*, 2012, **52**, (4), pp. 508–520
19 Shrivastava, V.K., Londhe, N.D., Sonawane, R.S., *et al.*: 'Computer-aided diagnosis of psoriasis skin images with HOS, texture and color features: a first comparative study of its kind', *Comput. Methods Programs Biomed.*, 2016, **126**, pp. 98–109
20 Maniruzzaman, M., Kumar, N., Abedin, M.M., *et al.*: 'Comparative approaches for classification of diabetes mellitus data: machine learning paradigm', *Comput. Methods Programs Biomed.*, 2017, **152**, pp. 23–34
21 Maniruzzaman, M., Rahman, M.J., Ahammed, B., *et al.*: 'Statistical characterization and classification of colon microarray gene expression data using multiple machine learning paradigms', *Comput. Methods Programs Biomed.*, 2019, **176**, pp. 173–193
22 Acharya, U.R., Sree, S.V., Krishnan, M.M.R., *et al.*: 'Automated classification of patients with coronary artery disease using grayscale features from left ventricle echocardiographic images', *Comput. Methods Programs Biomed.*, 2013, **112**, (3), pp. 624–632
23 Araki, T., Ikeda, N., Shukla, D., *et al.*: 'PCA-based polling strategy in machine learning framework for coronary artery disease risk assessment in intravascular ultrasound: a link between carotid and coronary grayscale plaque morphology', *Comput. Methods Programs Biomed.*, 2016, **128**, pp. 137–158
24 Martis, R.J., Acharya, U.R., Prasad, H., *et al.*: 'Application of higher order statistics for atrial arrhythmia classification', *Biomed. Signal Proc. Control*, 2013, **8**, (6), pp. 888–900
25 van Ravenswaaij-Arts, C.M., Kollee, L.A., Hopman, J.C., *et al.*: 'Heart rate variability', *Ann. Intern. Med.*, 1993, **118**, (6), pp. 436–447
26 Acharya, U.R., Mookiah, M.R.K., Sree, S.V., *et al.*: 'Atherosclerotic plaque tissue characterization in 2D ultrasound longitudinal carotid scans for automated classification: a paradigm for stroke risk assessment', *Med. Biol. Eng. Comput.*, 2013, **51**, (5), pp. 513–523
27 Saba, L., Jain, P.K., Suri, H.S., *et al.*: 'Plaque tissue morphology-based stroke risk stratification using carotid ultrasound: a polling-based PCA learning paradigm', *J. Med. Syst.*, 2017, **41**, (6), p. 98
28 Acharya, U.R., Faust, O., Alvin, A., *et al.*: 'Understanding symptomatology of atherosclerotic plaque by image-based tissue characterization', *Comput. Methods Programs Biomed.*, 2013, **110**, (1), pp. 66–75
29 Acharya, R.U., Faust, O., Alvin, A.P.C., *et al.*: 'Symptomatic vs. asymptomatic plaque classification in carotid ultrasound', *J. Med. Syst.*, 2012, **36**, (3), pp. 1861–1871
30 Suri, J.S.: 'Imaging based symptomatic classification and cardiovascular stroke risk score estimation', Google Patents, 2011
31 Acharya, U.R., Faust, O., Sree, S.V., *et al.*: 'An accurate and generalized approach to plaque characterization in 346 carotid ultrasound scans', *IEEE Trans. Instrum. Meas.*, 2011, **61**, (4), pp. 1045–1053
32 Saba, L., Ikeda, N., Deidda, M., *et al.*: 'Association of automated carotid IMT measurement and HbA1c in Japanese patients with coronary artery disease', *Diabetes Res. Clin. Pract.*, 2013, **100**, (3), pp. 348–353
33 Saba, L., Biswas, M., Kuppili, V., *et al.*: 'The present and future of deep learning in radiology', *Eur. J. Radiol.*, 2019, **114**, pp. 14–24
34 Biswas, M., Kuppili, V., Saba, L., *et al.*: 'State-of-the-art review on deep learning in medical imaging', *Front Biosci (Landmark Ed)*, 2019, **24**, pp. 392–426
35 Tandel, G.S., Balestrieri, A., Jujaray, T., *et al.*: 'Multiclass magnetic resonance imaging brain tumor classification using artificial intelligence paradigm', *Comput. Biol. Med.*, 2020, **122**, p. 103804
36 Shattuck, D.W., and Leahy, R.M.: 'Brainsuite: an automated cortical surface identification tool', *Med. Image Anal.*, 2002, **6**, (2), pp. 129–142
37 Rujirakul, K., So-In, C., and Arnonkijpanich, B.: 'PEM-PCA: a parallel expectation-maximization PCA face recognition architecture', *Scientific World J.*, 2014, **2014**, p. 468176
38 Golle, P.: 'Machine learning attacks against the Asirra CAPTCHA'. Proc. of the 15th ACM Conf. on Computer and communications security, Alexandria, VA, USA, 2008
39 Kaden, M., Riedel, M., and Hermann, W.: 'Border-sensitive learning in generalized learning vector quantization: an alternative to support vector machines', *Soft Comput.*, 2015, **19**, (9), pp. 2423–2434
40 Suk, H.-I., Lee, S.-W., Shen, D., *et al.*: 'Deep ensemble learning of sparse regression models for brain disease diagnosis', *Med. Image Anal.*, 2017, **37**, pp. 101–113
41 Abiwinanda, N., Hanif, M., Hesaputra, S.T., *et al.*: 'Brain tumor classification using convolutional neural network'. World Congress on Medical Physics and Biomedical Engineering 2018, Prague, Czech Republic (Springer, Singapore, 2019), pp. 183–189
42 Zhang, Y., Zhang, H., Chen, X., *et al.*: 'Strength and similarity guided group-level brain functional network construction for MCI diagnosis', *Pattern Recognit.*, 2019, **88**, pp. 421–430
43 Abrol, A., Bhattarai, M., Fedorov, A., *et al.*: 'Deep residual learning for neuroimaging: an application to predict progression to alzheimer's disease', *J. Neurosci. Methods*, 2020, **339**, p. 108701
44 Jing, R., Han, Y., Cheng, H., *et al.*: 'Altered large-scale functional brain networks in neurological wilson's disease', *Brain Imaging Behav.*, 2019, **14**, (5), pp. 1445–1455
45 Richhariya, B., Tanveer, M., Rashid, A.H., *et al.*: 'Diagnosis of Alzheimer's disease using universum support vector machine based recursive feature elimination (USVM-RFE)', *Biomed. Signal Proc. Control*, 2020, **59**, p. 101903
46 Liu, J., Pan, Y., Wu, F.-X., *et al.*: 'Enhancing the feature representation of multi-modal MRI data by combining multi-view information for MCI classification', *Neurocomputing*, 2020, **400**, pp. 322–332
47 Mirmehdi, M.: 'Handbook of texture analysis' (Imperial College Press, UK, 2008)
48 El-Baz, A., Gimel farb, G., and Suri, J.S.: 'Stochastic modeling for medical image analysis' (CRC Press, USA, 2015)
49 Biswas, M., Kuppili, V., Edla, D.R., *et al.*: 'Symtosis: a liver ultrasound tissue characterization and risk stratification in optimized deep learning paradigm', *Comput. Methods Programs Biomed.*, 2018, **155**, pp. 165–177
50 Biswas, M., Kuppili, V., Araki, T., *et al.*: 'Deep learning strategy for accurate carotid intima-media thickness measurement: an ultrasound study on Japanese diabetic cohort', *Comput. Biol. Med.*, 2018, **98**, pp. 100–117

1350911x, 2020, 25, Downloaded from https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/el.2020.2102 by CochraneItalia, Wiley Online Library on [26/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License