# AMSF: Attention-based multi-view slice fusion for early diagnosis of Alzheimer's disease

6 Yameng Zhang[1], Shaokang Peng [2], Zhihua Xue [3], Guohua Zhao [4], Qing Li [5,] Zhiyuan Zhu [6],
7 Yufei Gao [2,*] and Lingfei Kong [1,*]

9 [1] Department of Pathology, Henan Provincial People's Hospital & People's Hospital of
10 Zhengzhou University, Zhengzhou, Henan, China
11 [2] School of Cyber Science and Engineering, Zhengzhou University, SongShan Laboratory,
12 Zhengzhou, Henan, China
13 [3] Laboratory Animal Center, Academy of Medical Sciences, Zhengzhou University, Zhengzhou,
14 Henan, China
15 [4] Department of Magnetic Resonance Imageing, The First Affiliated Hospital of Zhengzhou
16 University, Zhengzhou, Henan, China
17 [5] State Key Lab of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing,
18 Beijing, China
19 [4] School of Communication and Information Engineering, Chongqing University of Posts and
20 Telecommunications, Chongqing, Chongqing, China

22 Corresponding Author:
23 Lingfei Kong[1,*]
24 No. 7, Weiwu Road, Zhengzhou, Henan, 450003, China
25 Email address: lfkong@zzu.edu.cn
26 Yufei Gao[2,*]
27 No. 97, Wenhua Road, Zhengzhou, Henan, 450003, China
28 Email address: yfgao@zzu.edu.cn

29

## Abstract

31 Alzheimer's disease (AD) is an irreversible neurodegenerative disease with a high prevalence in
32 the elderly population over 65 years of age. Intervention in the early stages of AD is of great
33 significance to alleviate the symptom. Recent advances in deep learning have shown extreme
34 advantages in computer-aided diagnosis of AD. However, most studies only focus on extracting
35 features from slices in specific direction or whole brain images, ignoring the complementarity
36 between features from different angles. To overcome the above problem, attention-based multi-
37 view slice fusion (AMSF) is proposed for accurate early diagnosis of AD. It adopts the fusion of
38 three-dimensional global features with multi-view 2D slice features by using an attention
39 mechanism to guide the fusion of slice features for each view, to generate a comprehensive

40    representation of the MRI images for classification. The experiments on the public dataset
41    demonstrate that AMSF achieves significant improvements over several previous promising
42    methods. It indicates that the better solution for AD early diagnosis is not only depends on the
43    large scale of dataset, but also the organic combination of feature construction strategy and deep
44    neural networks.

45    **Keywords**: Alzheimer's disease; magnetic resonance imageing; attention mechanism; multi-view
46    slice fusion

47

## Introduction

49    AD is a neurodegenerative disease with a high prevalence in people over 65 years of age
50    (Reiman et al. 2012). Previous studies have shown that the structural changes in the brain caused
51    by AD can be traced back 20 years before the onset of symptoms in patients (Barthelemy et al.
52    2020). In the early stages of AD, patients may not notice any significant changes in their brain
53    structure or activity, but some difficulties with memory recall or retention may appear. As AD
54    progresses, it leads to the formation of brain tissue lesions that impair and ultimately destroy
55    neurons responsible for various cognitive functions (Zarei et al. 2013), including deterioration of
56    memory and thinking skills, as well as a decline in physical abilities and independence. Patients
57    with AD may present symptoms such as memory loss, cognitive impairment, language
58    difficulties, and reduced mobility (Gaugler et al. 2022).
59        The global community is currently confronted with a significant demographic predicament
60    characterized by a rapid expansion in the population of older individuals. As per the United
61    Nations, the proportion of individuals aged 65 years and above in the overall global population is
62    projected to reach 9.7% by 2022, and further escalate to 16.4% by 2050 (ECONOMIC &
63    AFFAIRS. 2023). This unprecedented surge in ageing demographics presents formidable
64    challenges for healthcare systems, given the heightened vulnerability of older adults to chronic
65    and degenerative ailments. Among these conditions, AD stands out as a highly prevalent and
66    debilitating disorder that profoundly impacts the cognitive and functional capacities of countless
67    individuals across the globe.
68        Mild cognitive impairment (MCI) represents a pivotal transitional phase between normal
69    ageing and dementia, characterized by a discernible cognitive decline that does not significantly
70    impede daily functioning. MCI assumes a critical role as an early intervention window,
71    presenting a valuable opportunity to mitigate or forestall subsequent cognitive deterioration
72    (Wee et al. 2012). Extensive research has established that individuals diagnosed with MCI face a
73    heightened susceptibility to developing AD, with an annual conversion rate ranging from 10% to
74    15% (Roberts & Knopman 2013). Consequently, the implementation of timely and efficacious
75    medical interventions during the MCI stage holds the potential to safeguard neural cells against
76    further impairment and delay the onset of AD pathology, thereby contributing to a reduction in
77    the mortality associated with this incurable affliction (Odusami et al. 2022).
78        Neuroimageing serves as a valuable and indispensable tool in the clinical diagnosis of AD,
79    enabling the quantification of structural and functional alterations within the brain that
80    accompany disease progression. Among the diverse array of neuroimageing modalities available,
81    magnetic resonance imageing (MRI) has garnered considerable attention owing to its high spatial
82    resolution and non-invasive characteristics. Through MRI, intricate details regarding brain
83    volume, cortical thickness, white matter integrity, and cerebral blood flow in AD patients can be
84    gleaned. Notably, the advent of deep learning techniques has emerged as a formidable approach

85   for medical image analysis across a wide range of conditions, encompassing neurodegenerative
86   disorders, orthopaedic ailments, and cancer. Leverageing the intrinsic capacity to automatically
87   learn and extract intricate features from image data through the construction of multilayer neural
88   networks, deep learning methods transcend conventional machine learning approaches by
89   iteratively optimizing models with large-scale data. This obviates the need for manual feature
90   engineering, engenders enhanced diagnostic accuracy, and improves overall diagnostic
91   efficiency.
92       The integration of deep learning techniques into AD diagnostic research has been
93   instrumental in the development of algorithms aimed at supporting physicians in early diagnosis
94   and prognosis prediction. By harnessing the power of deep learning for MRI analysis, the
95   detection of AD at its nascent stages becomes attainable, thereby augmenting the diagnostic
96   capabilities and precision of healthcare professionals. This, in turn, facilitates the timely
97   implementation of intervention strategies to mitigate further cognitive decline. The exploration
98   of deep learning-based early AD diagnosis holds significant theoretical and practical
99   implications, encompassing the identification of initial brain alterations in AD patients, the
100  enhancement of AD diagnostic efficiency, the amelioration of the quality of life for individuals
101  afflicted by AD, and the advancement of deep learning theory as a whole.
102      Deep learning has become increasingly prominent in medical image analysis, surpassing
103  conventional machine learning algorithms in various domains (Lian et al. 2020). Notably, its
104  automated feature learning capability from raw data, without the need for human intervention or
105  domain expertise, distinguishes it as a highly advantageous approach. Among the array of deep
106  learning models, Convolutional Neural Networks (CNNs) have demonstrated remarkable success
107  and widespread adoption for medical image analysis. This can be attributed to their proficiency
108  in capturing both spatial and semantic information from images, thereby enabling robust and
109  accurate analysis in the medical field.
110      Korolev et al. (Korolev et al. 2017) explored the use of 3D CNNs for AD classification and
111  developed two 3D CNN models that achieved comparable results to traditional methods using
112  ADNI data. Cheng et al. (Cheng et al. 2017) employed 3D CNNs for AD classification, but they
113  amalgamated multiple 3D CNNs by training them on MRI data from distinct brain regions and
114  subsequently appending an FC layer to each one. Improved AD diagnosis performance of 3D
115  CNNs was achieved by Zhang et al. (Zhang et al. 2021) through the incorporation of an attention
116  mechanism, which enabled the network to selectively focus on relevant features. Spasov et al.
117  (Spasov et al. 2019) proposed a method to reduce the computational complexity of 3D CNNs by
118  using separable convolution techniques.
119      Pan et al. (Pan et al. 2022) invented an adaptive interpretable ensemble model
120  (3DCNN+EL+GA) that leverages the power of 3DCNN, ensemble learning and genetic
121  algorithm (GA) for AD classification and biomarker discovery. 246 base classifiers (3DCNN)
122  were trained on a dataset of 246 brain regions and a majority voting scheme    was employed to
123  select the optimal combination of base classifiers from the set of classifiers by using GA. Liu et
124  al. (Liu et al. 2020) developed a multi-task deep CNN model that performed both hippocampal
125  segmentation and disease classification tasks simultaneously. They combined a 3D densely
126  connected convolutional network (3D DenseNet) with the hippocampal segmentation results to
127  learn richer features for AD diagnosis.
128      Khvostikov et al. (Khvostikov et al. 2018) built 3D CNNs for AD classification by extracting
129  hippocampal ROIs from sMRI and DTI data. They also balanced the classes of different sizes by
130  using data augmentation methods and investigated the effect of ROI size on classification results.

131   Liu et al. (Liu et al. 2018) developed an end-to-end approach for AD classification by extracting
132   local fMRI image patches centred on predefined anatomical landmarks. These patches are
133   applied to capture both the local and global structural features from the images. However, many
134   of these 3D deep learning-based approaches still excessively rely on pre-determined ROIs before
135   the training of the network, which may limit the performance due to the presence of irrelevant
136   features in sMRI for AD diagnosis. Moreover, most of these studies only focus on binary
137   classification, which is not very helpful for determining the stage of the patient's situation.
138       Each individual's brain exhibits unique characteristics and may possess disease-related
139   features that cannot be fully captured by a single MRI slice. Consequently, the MRI slices of
140   patients may exhibit minimal deviation from those of healthy individuals, making classification
141   challenging. Previous studies have primarily focused on extracting features from specific slices
142   or the entire image, disregarding the features of slices from different views and the
143   complementary nature of features across these slices. Moreover, they have not effectively
144   utilized the comprehensive structural information available in whole-brain MRI scans (Lian et al.
145   2022). In light of this, Qiao et al. (Qiao et al. 2021) proposed a novel approach for early AD
146   diagnosis based on MRI, which involves extracting fused global features from multi-view slice
147   features. They employed a simple splicing technique to combine the features of multiple slices
148   from the same view. However, it is restricted by an assumption that all slices are of equal
149   importance for the classification task. Different slices may contribute differently to disease
150   features. Therefore, using equal weights for feature fusion may not effectively capture the
151   relevant features, potentially resulting in lower classification accuracy.
152       Actually, the slices from the same view capture diverse brain regions and exhibit distinct
153   features, essentially representing channel-specific mappings that reflect varying degrees of
154   importance in the slice clusters. Additionally, it is essential to recognize that slices from different
155   locations are not isolated entities but interconnected, collectively forming a comprehensive
156   feature representation of the slice cluster in that specific direction. Hence, when fusing the
157   features of different slices from the same view, it becomes crucial to consider both the
158   significance of the information carried by each slice and the contextual relationship that exists
159   between them. By incorporating these factors into the fusion process, a more robust and
160   informative representation can be achieved, facilitating improved accuracy in capturing essential
161   features for classification tasks, such as early AD diagnosis based on MRI data.
162       In this dissertation, a novel approach to the early diagnosis of AD called ASMF is proposed.
163   Firstly, the Multi-view Slice-level Feature Extraction (MSFE) method is employed to acquire
164   slice-level features from three distinct views (sagittal, coronal, and cross-sectional) by repeatedly
165   slicing the 3D MRI and leverage three separate 2D sub-networks to extract features from each
166   view. Then, an attention mechanism is incorporated to guide the fusion of slice features for each
167   view, assigning varying weights to individual slices based on their respective importance.
168   Secondly, global features are extracted from the entire MRI images by using a 3D CNN to
169   complement the slice-level features. Finally, the slice-level and global features are fused to
170   generate a more comprehensive feature representation for the classifier. The key contributions of
171   this study can be concluded as follows:
172   1. To address the limitations of relying on feature extraction in a specific direction, this study
173   proposes a novel MSFE-based approach for early diagnosis of AD.
174   2. This is the first study that incorporates the self-attention mechanism and the fusion of multi-
175   view features to construct a comprehensive representation of the MRI images for classification.

176  3. According to the experimental results, the proposed method outperforms other recently
177  published promising approaches.
178      The remainder of this paper is arranged as follows. The materials and methods are described
179  in Section 2. Section 3 provides experimental results and corresponding discussion. At last, the
180  summary of this study is given in Section 4.
181
182

## Materials & Methods

183

184  In contrast to previous studies that typically rely on a single slice from a specific view for AD
185  diagnosis, we utilize multiple slices from three views of 3D MRI scans to extract features. As
186  shown in Fig. 1, by incorporating information from various slices, we aim to capture the full
187  extent of brain damage caused by AD, accounting for patient heterogeneity. To guide the fusion
188  of slice features, an attention mechanism is employed that assigns different weights to slices
189  based on their relevance for classification. It ensures that the most informative slices contribute
190  more significantly to the diagnostic process. Furthermore, we integrate the slice features from the
191  three views with global features extracted from the entire MRI images, resulting in a feature
192  representation that comprehensively reflects the overall brain state of the patient.

### Multi-view slicing feature extraction

193

194  To facilitate the analysis of the 3D MRI data, we initially partitioned it into three distinct planes:
195  sagittal, coronal, and transverse planes. Each plane represents a different orientation of the brain
196  and provides unique information about its structural characteristics. To ensure a comprehensive
197  assessment, we extracted a total of 40 slices per view, thereby constructing a robust slice cluster.
198  Figure 2 visually illustrates the resulting set of images, showcasing the diversity and coverage
199  achieved across the different planes. By encompassing multiple slices from each view, our
200  approach captures a broader range of relevant features, enabling a more thorough examination of
201  the brain's structural attributes.

202      To extract features from each slice of the 3D MRI data, we used a slice-level feature
203  extraction network that takes slice clusters as input. A slice cluster consists of 40 slices from one
204  of three possible views: sagittal (x), coronal (y) or transverse (z), which represent different
205  orientations of the brain structure and contain different types of features. Therefore, we designed
206  a separate Slice Feature Extraction Network (SFEN) for each view. For instance, the sagittal
207  view (x), $C_x$ denotes the cluster of slices in this direction. Each slice in this cluster has an index i
208  that ranges from 1 to 40. Thus,   $C_x$ can be written as

209
$$C_x = \left[ C_x^1, C_x^2, ..., C_x^n \right] \qquad (1)$$

210  where n denotes the number of slices in the x-direction. The feature extraction of the i-th slice in
211  the x-direction can be expressed as

212
$$T_x^i = AvgPool(F_x(C_x^i, W_x^i)) \qquad (2)$$

213  where $F_x$ represents the feature extraction function consisting of multiple blocks containing a
214  3*3 convolutional layer, BN layer, activation function ReLU and maximum pooling layer, as
215  shown in Fig. 3. Besides, $W_x^i$ denotes the convolutional layer weight of the i-th slice in the x-
216  direction, $T_x$ stands for the features of the cluster of slices in the x-direction after $F_x$. Then the
217  slice cluster feature in the x-direction can be expressed as

218 
$$T_{\mathrm{x}} = \left[ T_x^1, T_x^2, ..., T_x^n \right] \qquad (3)$$

219

## Attention-based slice feature fusion

221 To integrate the distinct features within each view (x, y, or z), we leverage the notion of a "slice
222 cluster" comprising 40 slices that possess unique characteristics. These features serve as
223 mappings for specific brain structures observed from their respective perspectives. Notably,
224 these features are not isolated entities, but rather interconnected across various locations within
225 the slice cluster. Consequently, they collectively establish a comprehensive feature
226 representation for the given view. To effectively merge these features, it is imperative to consider
227 both their significance and their interdependencies. To address this, we propose a novel
228 mechanism termed Slices Fusion Attention (SFA). SFA employs self-attention to capture
229 contextual information among the slices and assigns attention weights to each slice based on its
230 relative importance and contribution to the overall feature representation of the view. By
231 incorporating this attention-based weighting scheme, SFA effectively balances the significance
232 of different slices while enriching them with contextual information derived from their
233 interrelationships. Figure 4 provides an illustrative depiction of the structural composition of
234 SFA.
235     We feed the slice cluster feature $T_x$ into SFA as input. $T_x$ has a dimension of 40*1*128 and it
236 is obtained by concatenating the features of 40 slices along the channel dimension after applying
237 SFEN. To reduce the number of channels from 128 to 1, we use a 1*1 convolution layer that
238 compresses $T_x$ into a single-channel feature map. This gives us an aggregated feature that
239 represents the fusion of 40 slices. Next, we apply a Softmax function to $T_x$ and multiply it with
240 $S_x$. This way, we obtain $S_x$ that contains contextual information between slices weighted by their
241 attention scores. We can write this process as
242 
$$S_x = T_x^* * Softmax(Conv(T_x)) \qquad (4)$$

243 where Conv means 1*1 convolution operation, and $T_x^*$ is obtained by reducing one dimension of
244 $T_x$ .
245     $S_x$ contains the contextual relationship between different slices, which needs to be assigned to
246 each slice by calculating the slice feature weights of different channels. Firstly, $S_x$ is expanded
247 by one dimension and the number of channels is reduced by 1*1 convolution, then BatchNorm
248 and ReLU activation function operations are performed, and the number of channels is raised to
249 the original number by 1*1 convolution, denoted as

250

251 
$$A_x = Conv(BR(Conv(S_x^{'}))) \qquad (5)$$

252 where $S_x^{'}$ is obtained by adding one dimension to $S_x$, and BR denotes the BatchNorm and ReLU
253 activation functions.
254     $A_x$ is applied to represent the reweighted channel features that reflect how much each slice
255 contributes to $S_x$. To obtain the slice fusion feature that incorporates both channel weights and
256 contextual relationships between slices, we multiply $A_x$ with $T_x$ along the channel axis and sum
257 them up. We denote this final output as
258 
$$F_x = \sum_{i=1}^{40} T_x^i * A_x^i \qquad (6)$$

259     where 40 denotes the number of slices and $A_x^i$ denotes the i-th channel of weight $A_x$.

260

261     **Global feature extraction**

262     The global feature extraction (GFE) component contains four blocks, each encompassing a

263     sequence of operations: a 3D convolutional layer, a 3D batch normalization layer (BN), a

264     rectified linear unit (ReLU) activation function, and a 3D maximum pooling layer. Subsequently,

265     a 3D average pooling layer is employed to convert the multichannel features into a vector that

266     encapsulates the global information. Figure 5 provides the structure of GFE. Following the

267     acquisition of the global features, adaptive averageing pooling is applied to generate a one-

268     dimensional vector. The multi-view slice-level features are then concatenated with the global

269     vector. Finally, a fully connected layer is employed to obtain the ultimate classification

270     outcomes.

271

272     **Dataset and processing**

273     The Alzheimer's Disease

（1.项目研究的目的、意义，突出说明项目实施后对郑州市产业发展的预期作用；2.项目已有技术基础和专利储备描述，包括：本项目研究的关键技术及难点，现有技术水平，国内外相关的研究开发情况及知识产权现状，有关的主要论文、专著情况，项目技术实现主要面临的风险和应对措施，以往承担国家、省级等各类科技计划项目完成情况等）

1. 项目研究目的、意义

2. 项目已有技术基础和专利储备

2.1 本项目研究的关键技术及难点

漏洞挖掘技术

敏感数据加密技术

动态防御决策技术

2.2 现有技术水平

2.3 国内外相关的研究开发情况

274 Neuroimageing Initiative (ADNI) dataset (http://adni.loni.usc.edu/) is employed in this study.
275 ADNI provides data processed by standard volumetric analysis methods, including gradient non-
276 linearity correction, B1 correction, N3 correction, CAT12 for extraneous tissue removal,

277 alignment and smoothing operations. Then, a total of 351 3D-MRI scans for NC subjects are
278 acquired (301 3D-MRI scans for AD subjects and 331 3D-MRI scans for MCI subjects).
279     In the experiments, the ADNI data were utilized and underwent several preprocessing steps
280 before being fed into the feature extraction network, as depicted in Fig. 4. Initially, the
281 background information that is unrelated to the classification task was eliminated. Subsequently,
282 the image size was adjusted to 90*90*90, and the image density was normalized based on the
283 mean and standard values of the non-zero region. The 3D MRI data were then sliced according
284 to three directions, and for each view, the middle 40 slices were selected. Finally, feature
285 extraction operations were conducted on the obtained slices.
286

## Results and Discussion

### Ablation experiments

289 To investigate the performance of the proposed framework, the indicators of Acc (Accuracy),
290 Sen (Sensitivity), Spe (Specificity), Pre (Precision) and F1 (F1-score) are employed for
291 evaluation. The ablation experiments are conducted with the detailed definition as follows:

292 1. 3D: 3D features are only obtained by 3D GFE.
293 2. 2D: only employ 2D features learned by MSFE.
294 3. 2D+3D: 2D and 3D features are extracted by MSFE and GFE.
295 4. SE+2D+3D: the combination of 2D+3D with SE.
296 5. AMSF (SFA+2D+3D): SFA-guided 2D+3D.

297     The results of the ablation experiments are shown in Table 1. It can be seen that the
298 classification accuracy of 3D is the lowest (76.9%). The 2D method provides better performance
299 with the classification accuracy improvement to 89.4%, which is 12.5% higher than 3D. With the
300 help of fused global features, 2D+3D achieves 91.6% Acc better than only using 2D or 3D
301 strategy. However, the inclusion of SE results in a decrease in classification accuracy by 1.8%,
302 suggesting that SE fails to effectively address the imbalanced importance among different slices
303 within the same view for this particular task. At last, it is demonstrated that AMSF reaches the
304 highest accuracy of 94.3%, surpassing that of the 2D+ 3D and SE+2D+3D by 2.7% and 4.5%
305 respectively. It also exceeds in F1 scores, Sen, Spe and Pre. This indicates that SFA effectively
306 integrates contextual relationships between different slices, enabling it to balance their
307 importance within a given view.

308     The curve of validation loss for each epoch is shown in Fig. 6. It can be seen that during the
309 training process, the training loss decreases continuously in the first 20 epochs and keeps stable.
310 The validation loss declines in the first 10 epochs, then falls in fluctuation before 25 epochs, and
311 finally stabilises after 35 epochs. Figure 7 gives the curve of the accuracy for each epoch. With a
312 zigzag rise before 35 epochs, the validation accuracy reaches saturation. As can be seen in Fig. 8,
313 the classification of AD achieves the highest accuracy (97.0%), indicating that ASMF is more
314 sensitive to the features of AD. The classification accuracy for NC is 95.4% with 4.6%
315 misclassifying NC to MCI. For MCI, it obtains the worst performance (90.4%), as well as an
316 8.2% misclassifying MCI to AD. It can be observed that the subtle difference in features between
317 MCI and AD are formidable to extract, which is still a significant problem in AD diagnosis.

### Comparison with other methods

319 An experimental comparison with previous promising approaches is also arranged, including two
320 traditional machine learning-based approaches (JMMP-LRR, Liner SVM) and three deep
321 learning-based methods (DemNet, Automatic CL and AdaBoost). The abstracts of these works
322 are listed as follows:

323 1. Automatic CL (Gracias & Silveira 2022): curriculum learning is employed in the early
324    diagnosis of AD based on a 3D CNN network, by incorporating task complexity, cognitive
325    test scores, and ROI features, thereby enhancing the accurate classification of MCI.

326 2. Liner SVM (Yuan et al. 2022): Mr cortical and ApoE4 gene features are explored for AD
327    classification and the optimal performance is achieved by an SVM classifier with higher
328    sensitivity and specificity.

329 3. JMMP-LRR (Sheng et al. 2020): it aims to better alleviate the problem of small-sample, ul-
330    tra-high-dimensional features, accompany by stable AD classification accuracy.

331 4. DemNet (Billones et al. 2016): an improved 16-layer VGGNet architecture is proposed with
332    SOTA (state-of-the-art) classification results of AD vs. MCI vs. NC.

333 5. AdaBoost (Buyrukoğlu 2021): an ensemble learning method is designed for AD diagnosis
334    with SOTA performance compared with different ensemble learning methods.

335 Table 2 shows the classification results on the ADNI dataset compared with other previous
336 promising methods. It can be seen that ASMF achieves the best performance with 94.3%
337 classification accuracy, which is 1.6% -7.1% higher than other related works. Surprisingly, only
338 by using a small scale of dataset, traditional machine learning-based methods (Linner SVM and
339 AdaBoost) surpass the other three deep learning-based approaches. Especially for Linner SVM,
340 it outperforms Automatic CL and DemNet without the help of DNNs. Together these results
341 provide important insights into the better solution for AD early diagnosis not only depends on
342 the large scale of dataset, but also the organic combination of feature construction strategy and
343 deep neural networks.

344

## Conclusions

346 This study proposes a novel AD diagnosis approach called AMSF to address the limitation of
347 relying on the feature extraction in a specific direction. It incorporates multiple slices with
348 feature extraction from different views of 3D MRI images and the fusion of slice cluster features
349 from each view achieved through an attention mechanism. Furthermore, the GFE by the 3D
350 network fusion is combined to obtain a comprehensive MRI feature representation. Based on the
351 AD vs. MCI vs. NC classification tasks on the public ADNI dataset, the experimental results
352 demonstrate significant advantages over several SOTA methods. Although the promising
353 experimental results achieves, the drawback still exists to be addressed. For example, a dataset
354 with a limited size may restrict the generalization of the proposed model for practical
355 applications. In future, diverse datasets with large scales are required to enhance the robustness
356 and reliability of the method.

357

358

359 **Acknowledgments**

360 **Author Contributions**

361 S.P. and Y.G provided the idea and experiments; Y.Z, S.P. and Y.G. wrote the contents of the

362 article; Y.Z., Z.X., G.Z., Q.L, Z.Z reviewed and edited the manuscript, L.K and Y.G supported

363 the study. All authors have read and agreed to the published version of the manuscript.

364 **Data Availability Statement**

365 In this study, we used a publicly available MRI dataset: ADNI (http://adni.loni.usc.edu/).

366 **Acknowledgments**

367 This work was supported in part by the Nature Science Foundation of China (62006210,

368 62001284, 62206252), the Key Scientific and Technology Project in Henan Province of China

369 (221100210100), the Key Project of Collaborative Innovation in Nanyang (22XTCX12001), the

370 Research Foundation for Advanced Talents of Zhengzhou University (32340306).

371 **Conflicts of Interest**

372 The authors declare no conflict of interest.

373

374

375 **References**

376

377 Barthelemy NR, Li Y, Joseph-Mathurin N, Gordon BA, Hassenstab J, Benzinger TLS, Buckles
378       V, Fagan AM, Perrin RJ, Goate AM, Morris JC, Karch CM, Xiong C, Allegri R, Mendez
379       PC, Berman SB, Ikeuchi T, Mori H, Shimada H, Shoji M, Suzuki K, Noble J, Farlow M,
380       Chhatwal J, Graff-Radford NR, Salloway S, Schofield PR, Masters CL, Martins RN,
381       O'Connor A, Fox NC, Levin J, Jucker M, Gabelle A, Lehmann S, Sato C, Bateman RJ,
382       McDade E, and Dominantly Inherited Alzheimer N. 2020. A soluble phosphorylated tau
383       signature links tau, amyloid and the evolution of stages of dominantly inherited
384       Alzheimer's disease. *Nat Med* 26:398-407. 10.1038/s41591-020-0781-z

385 Billones CD, Demetria OJLD, Hostallero DED, and Naval PC. 2016. DemNet: A Convolutional
386       Neural Network for the detection of Alzheimer's Disease and Mild Cognitive Impairment.
387       2016 IEEE Region 10 Conference (TENCON). p 3724-3727.

388 Buyrukoğlu S. 2021. Improvement of Machine Learning Models' Performances based on
389       Ensemble Learning for the detection of Alzheimer Disease. 2021 6th International
390       Conference on Computer Science and Engineering (UBMK). p 102-106.

391 Cheng D, Liu M, Fu J, and Wang Y. 2017. Classification of MR brain images by combination of
392       multi-CNNs for AD diagnosis. Ninth international conference on digital image processing
393       (ICDIP 2017): SPIE. p 875-879.

394 ECONOMIC UNDF, and AFFAIRS. S. 2023. *World population prospects 2022: Summary of*
395       *results*: UN.

396 Gaugler J, James B, Johnson T, Reimer J, Solis M, Weuve J, Buckley RF, and Hohman TJ.
397       2022. 2022 Alzheimer's disease facts and figures. *Alzheimers & Dementia* 18:700-789.

398 Gracias C, and Silveira M. 2022. Curriculum learning for early Alzheimer's Disease diagnosis.
399       *Annu Int Conf IEEE Eng Med Biol Soc* 2022:4777-4780.
400       10.1109/EMBC48229.2022.9871601

401 Khvostikov A, Aderghal K, Benois-Pineau J, Krylov A, and Catheline G. 2018. 3D CNN-based
402       classification using sMRI and MD-DTI images for Alzheimer disease studies. *arXiv*
403       *preprint arXiv:180105968*.

404 Korolev S, Safiullin A, Belyaev M, and Dodonova Y. 2017. Residual and plain convolutional
405     neural networks for 3D brain MRI classification. 2017 IEEE 14th international symposium
406     on biomedical imageing (ISBI 2017): IEEE. p 835-838.
407 Lian C, Liu M, Pan Y, and Shen D. 2022. Attention-Guided Hybrid Network for Dementia
408     Diagnosis With Structural MR Images. *IEEE Trans Cybern* 52:1992-2003.
409     10.1109/TCYB.2020.3005859
410 Lian C, Liu M, Zhang J, and Shen D. 2020. Hierarchical Fully Convolutional Network for Joint
411     Atrophy Localization and Alzheimer's Disease Diagnosis Using Structural MRI. *IEEE*
412     *Trans Pattern Anal Mach Intell* 42:880-893. 10.1109/TPAMI.2018.2889096
413 Liu M, Li F, Yan H, Wang K, Ma Y, Alzheimer's Disease Neuroimageing I, Shen L, and Xu M.
414     2020. A multi-model deep convolutional neural network for automatic hippocampus
415     segmentation and classification in Alzheimer's disease. *Neuroimage* 208:116459.
416     10.1016/j.neuroimage.2019.116459
417 Liu M, Zhang J, Adeli E, and Shen D. 2018. Landmark-based deep multi-instance learning for
418     brain disease diagnosis. *Med Image Anal* 43:157-168. 10.1016/j.media.2017.10.005
419 Odusami M, Maskeliunas R, and Damasevicius R. 2022. An Intelligent System for Early
420     Recognition of Alzheimer's Disease Using Neuroimageing. *Sensors (Basel)* 22.
421     10.3390/s22030740
422 Pan D, Luo G, Zeng A, Zou C, Liang H, Wang J, Zhang T, Yang B, and Initiative AsDN. 2022.
423     Adaptive 3DCNN-Based Interpretable Ensemble Model for Early Diagnosis of
424     Alzheimer's Disease. *IEEE Transactions on Computational Social Systems*.
425 Qiao H, Chen L, and Zhu F. 2021. A Fusion of Multi-view 2D and 3D Convolution Neural
426     Network based MRI for Alzheimer's Disease Diagnosis. *Annu Int Conf IEEE Eng Med*
427     *Biol Soc* 2021:3317-3321. 10.1109/EMBC46164.2021.9629923
428 Reiman EM, Quiroz YT, Fleisher AS, Chen K, Velez-Pardo C, Jimenez-Del-Rio M, Fagan AM,
429     Shah AR, Alvarez S, Arbelaez A, Giraldo M, Acosta-Baena N, Sperling RA, Dickerson B,
430     Stern CE, Tirado V, Munoz C, Reiman RA, Huentelman MJ, Alexander GE, Langbaum
431     JB, Kosik KS, Tariot PN, and Lopera F. 2012. Brain imageing and fluid biomarker
432     analysis in young adults at genetic risk for autosomal dominant Alzheimer's disease in
433     the presenilin 1 E280A kindred: a case-control study. *Lancet Neurol* 11:1048-1056.
434     10.1016/S1474-4422(12)70228-4
435 Roberts R, and Knopman DS. 2013. Classification and epidemiology of MCI. *Clin Geriatr Med*
436     29:753-772. 10.1016/j.cger.2013.07.003
437 Sheng J, Shao M, Zhang Q, Zhou R, Wang L, and Xin Y. 2020. Alzheimer's disease, mild
438     cognitive impairment, and normal ageing distinguished by multi-modal parcellation and
439     machine learning. *Sci Rep* 10:5475. 10.1038/s41598-020-62378-0
440 Spasov S, Passamonti L, Duggento A, Lio P, Toschi N, and Alzheimer's Disease
441     Neuroimageing I. 2019. A parameter-efficient deep learning approach to predict
442     conversion from mild cognitive impairment to Alzheimer's disease. *Neuroimage* 189:276-
443     287. 10.1016/j.neuroimage.2019.01.031
444 Wee CY, Yap PT, Zhang D, Denny K, Browndyke JN, Potter GG, Welsh-Bohmer KA, Wang L,
445     and Shen D. 2012. Identification of MCI individuals using structural and functional
446     connectivity networks. *Neuroimage* 59:2045-2056. 10.1016/j.neuroimage.2011.10.015
447 Yuan Z, Yao X, and Bu X. 2022. Classification of Alzheimer's Disease Using Conventional
448     Machine Learning Methods with Cortical and Genetic Characteristics. 2022 IEEE 2nd
449     International Conference on Power, Electronics and Computer Applications (ICPECA). p
450     303-306.
451 Zarei M, Beckmann CF, Binnewijzend MA, Schoonheim MM, Oghabian MA, Sanz-Arigita EJ,
452     Scheltens P, Matthews PM, and Barkhof F. 2013. Functional segmentation of the
453     hippocampus in the healthy human brain and in Alzheimer's disease. *Neuroimage* 66:28-
454     35. 10.1016/j.neuroimage.2012.10.071

455 Zhang J, Zheng B, Gao A, Feng X, Liang D, and Long X. 2021. A 3D densely connected
456       convolution neural network with connection-wise attention mechanism for Alzheimer's
457       disease classification. *Magn Reson Imageing* 78:119-126. 10.1016/j.mri.2021.02.001
458

# Figure 1

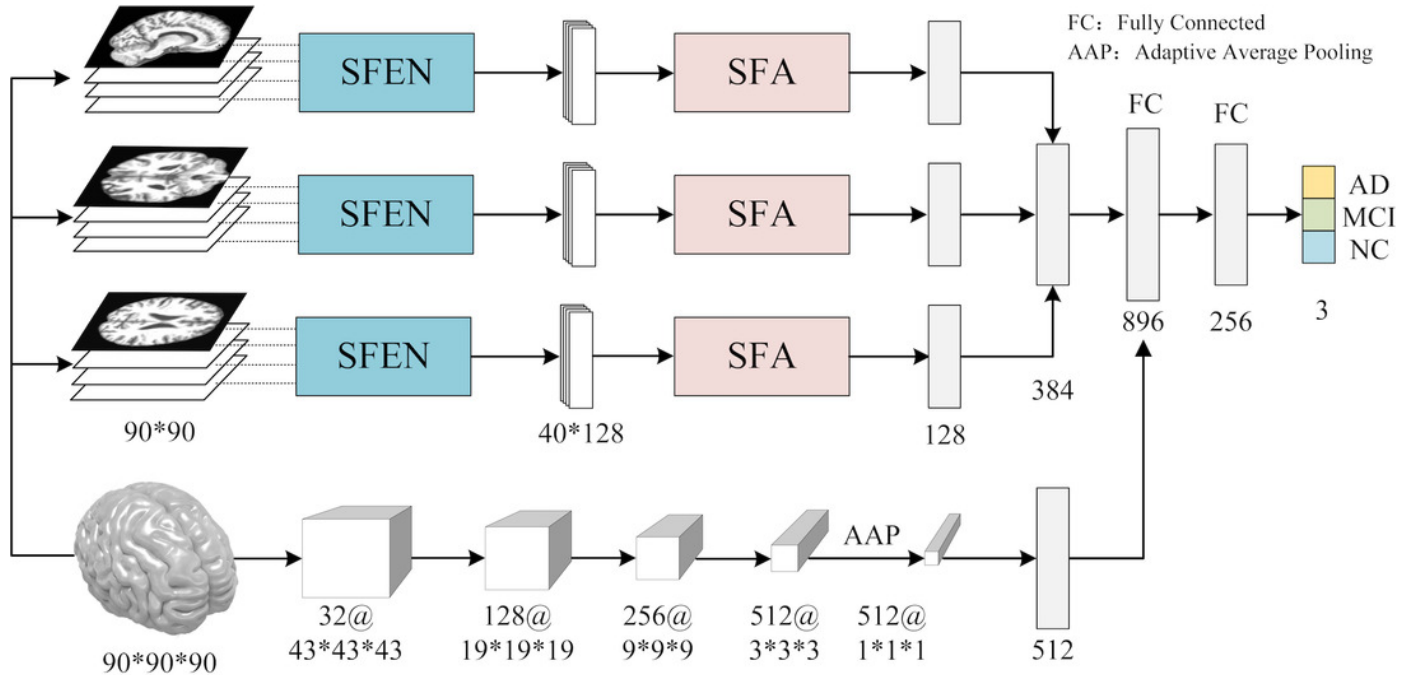
Figure 1. The framework of AMSF.

# Figure 2

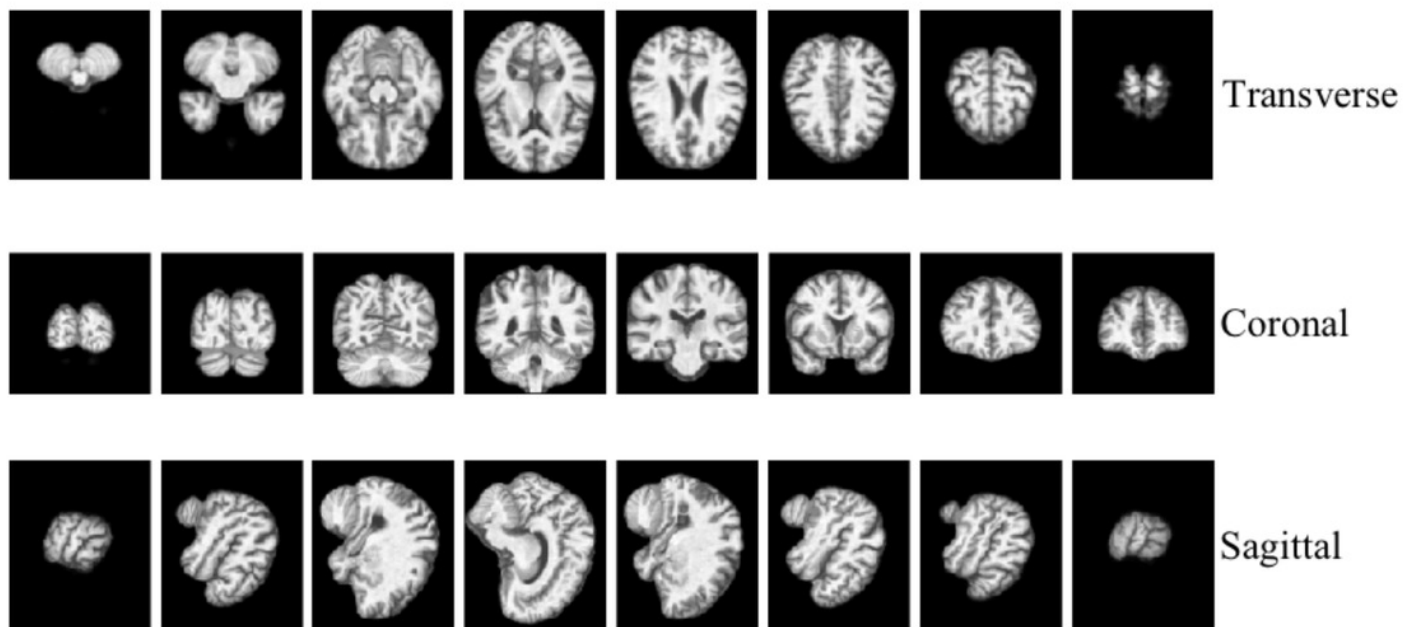
Figure 2. 3D MRI data slicing in three directions.

# Figure 3

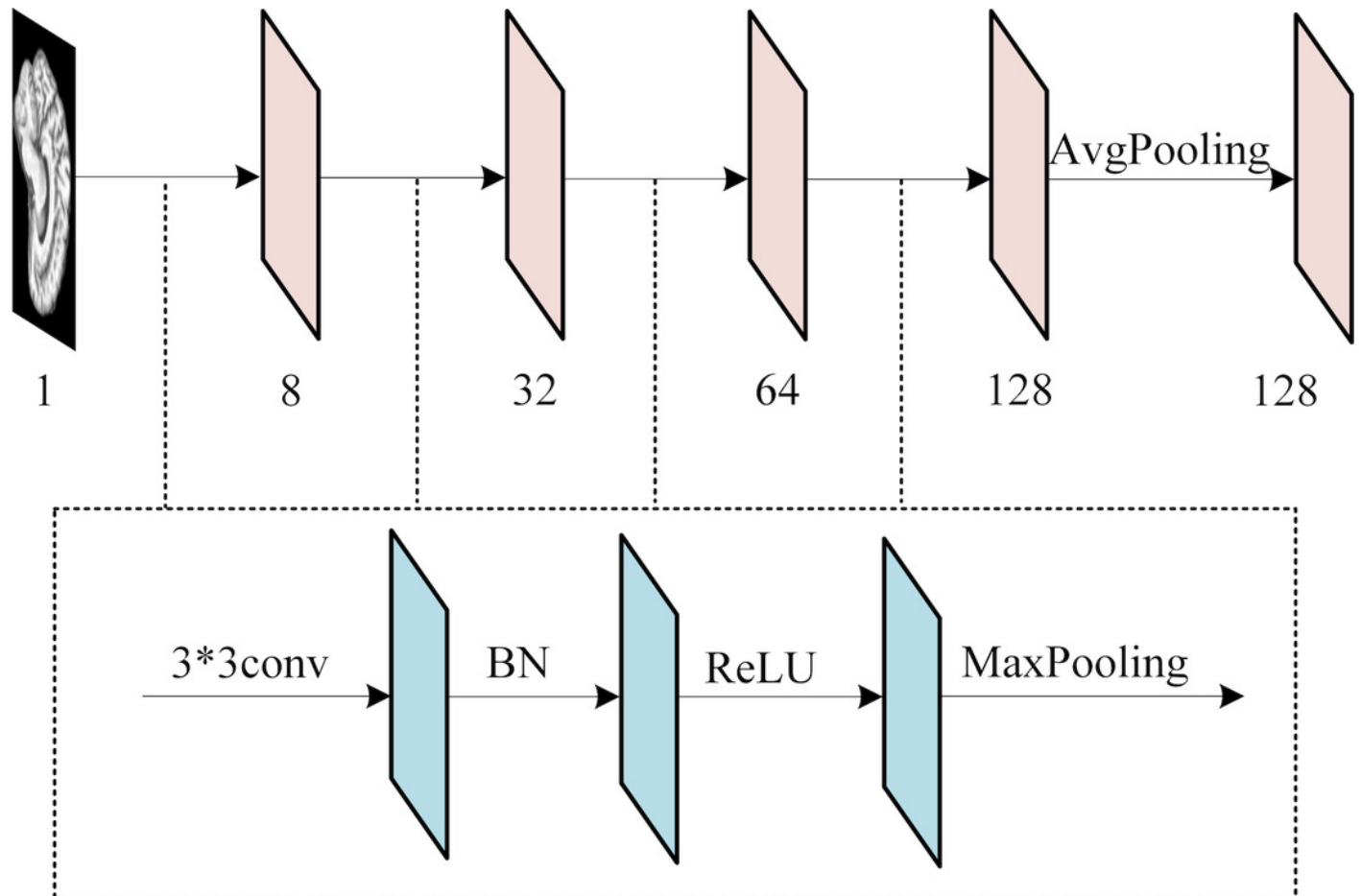
Figure 3. The architecture of SFE.

# Figure 4

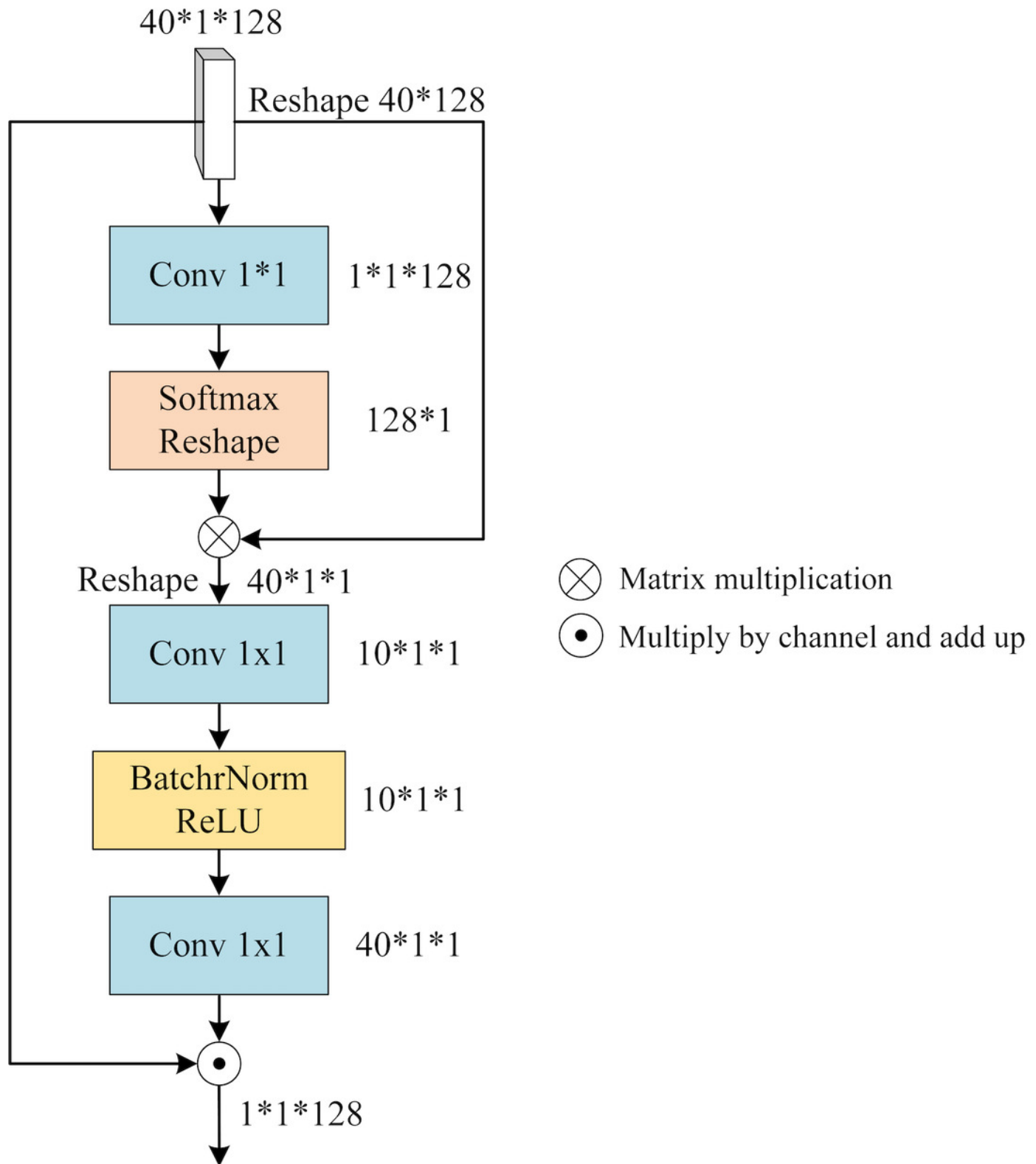
Figure 4. The architecture of SFA.

# Figure 5

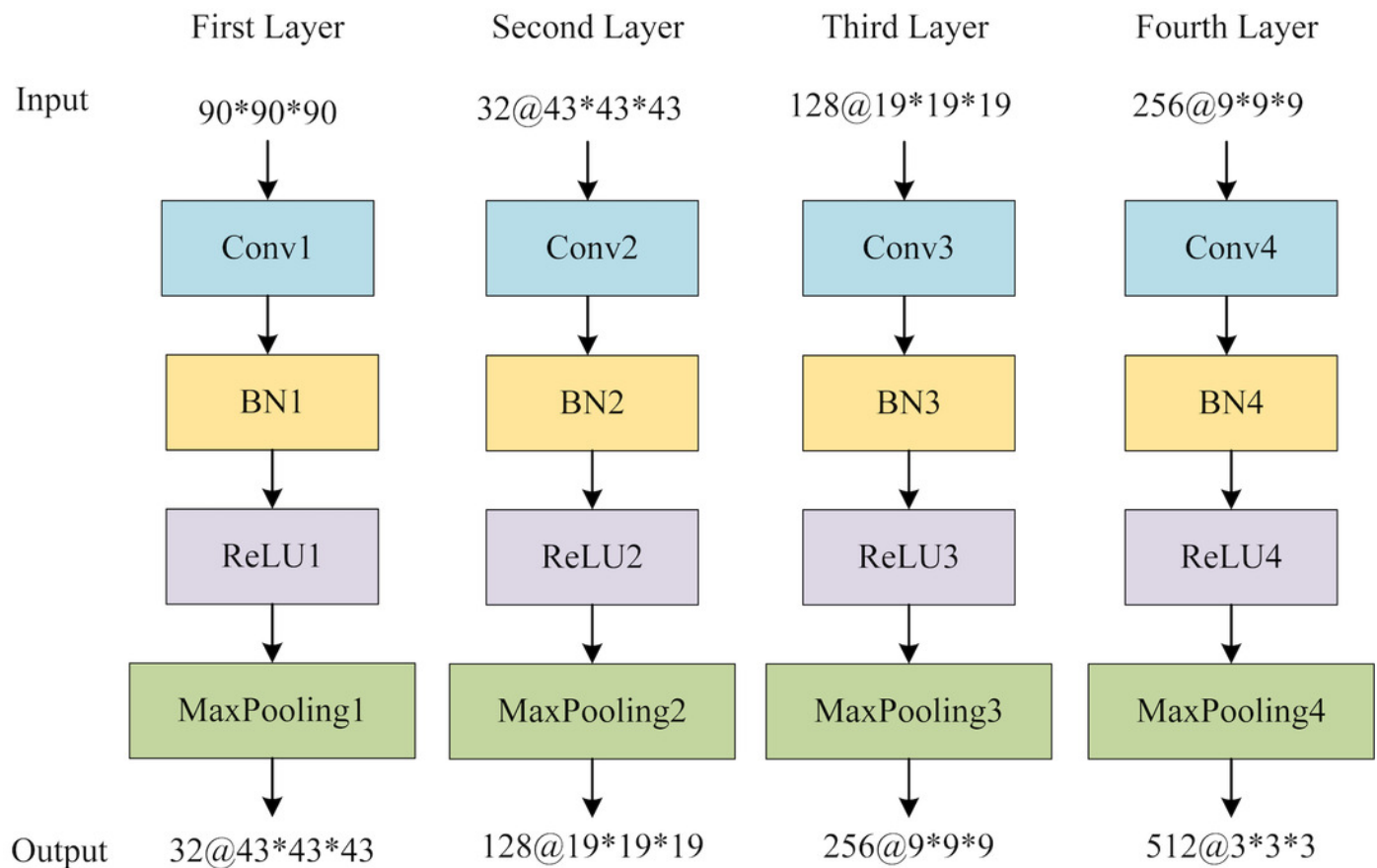
Figure 5. The structure of GFE network.

# Figure 6

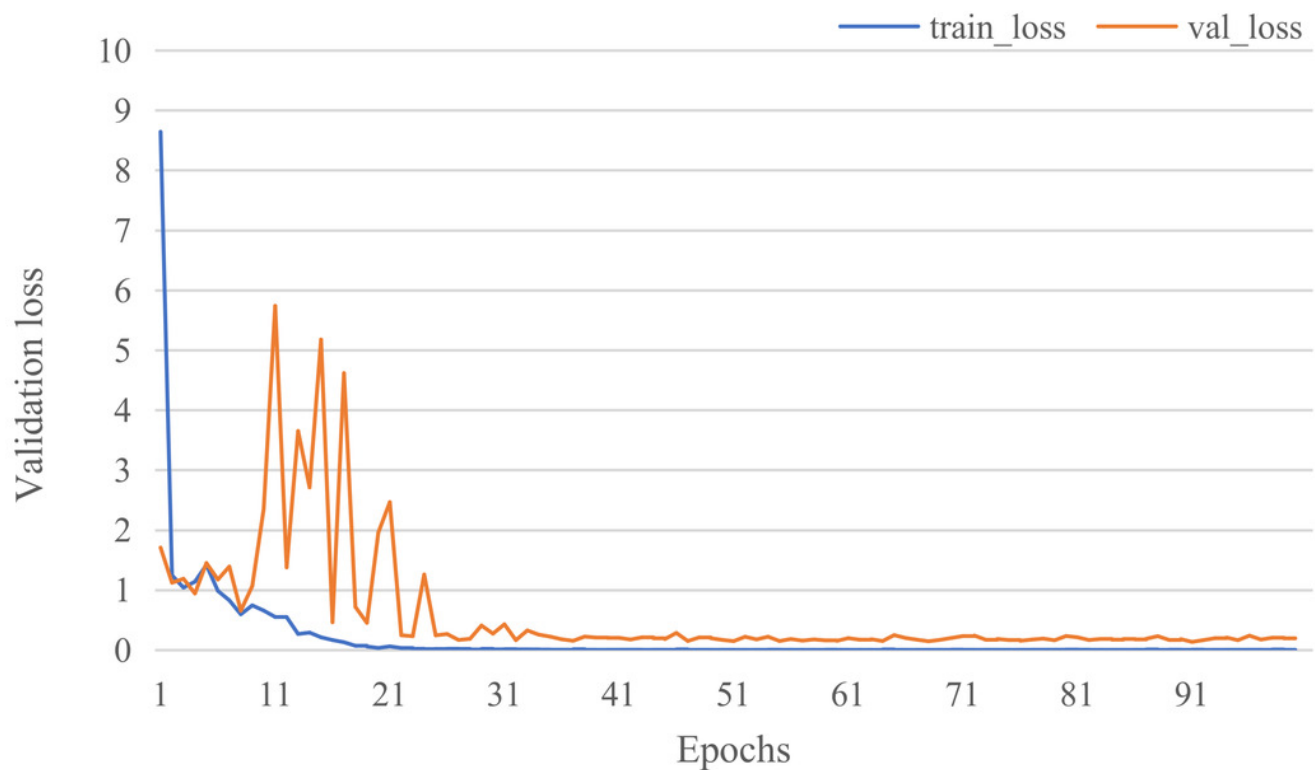
Figure 6. The curve of validation loss of AMSF.

# Figure 7

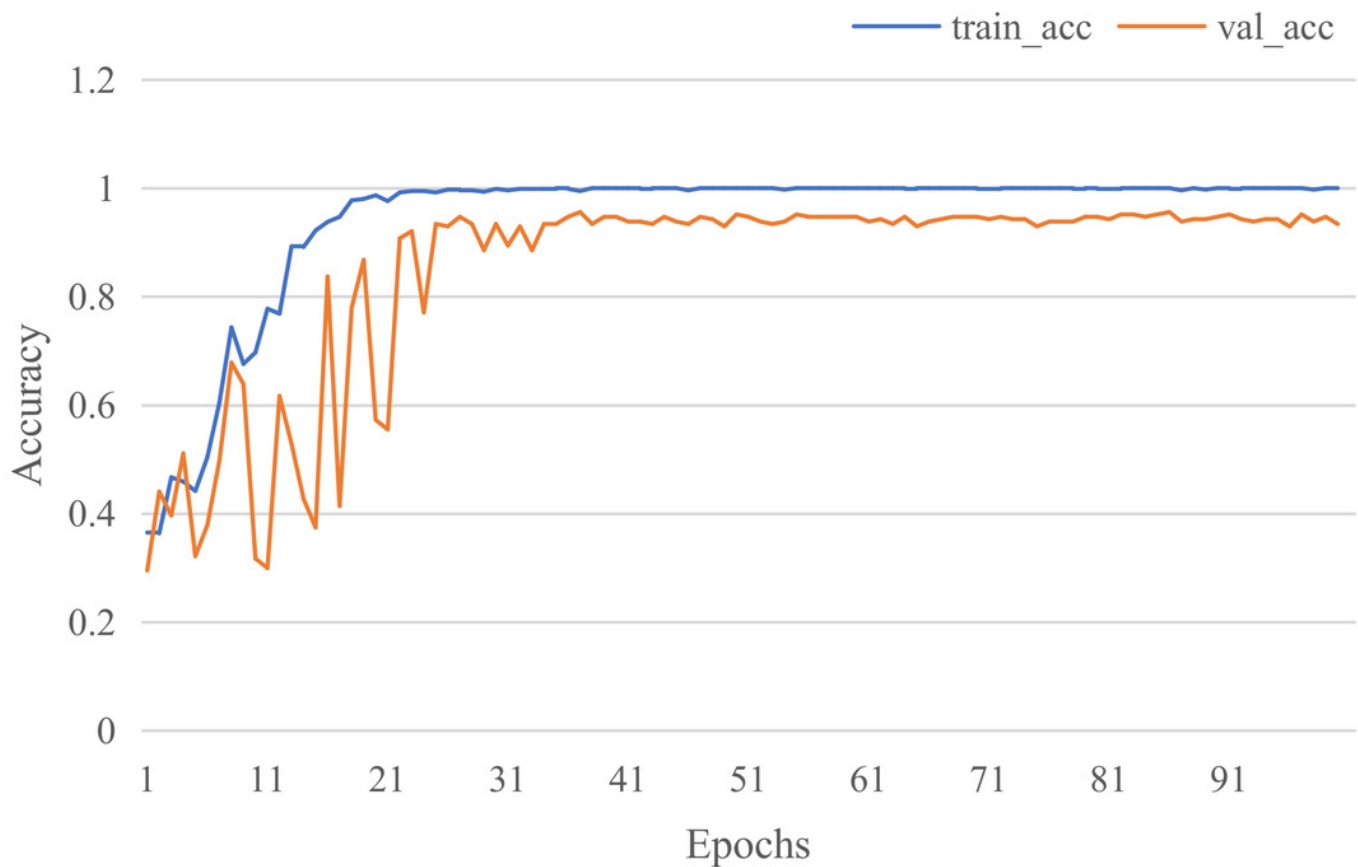
Figure 7. The curve of accruracy of AMSF.

# Figure 8

Figure 8. Confusion matrix of of AMSF in ablation experiment.

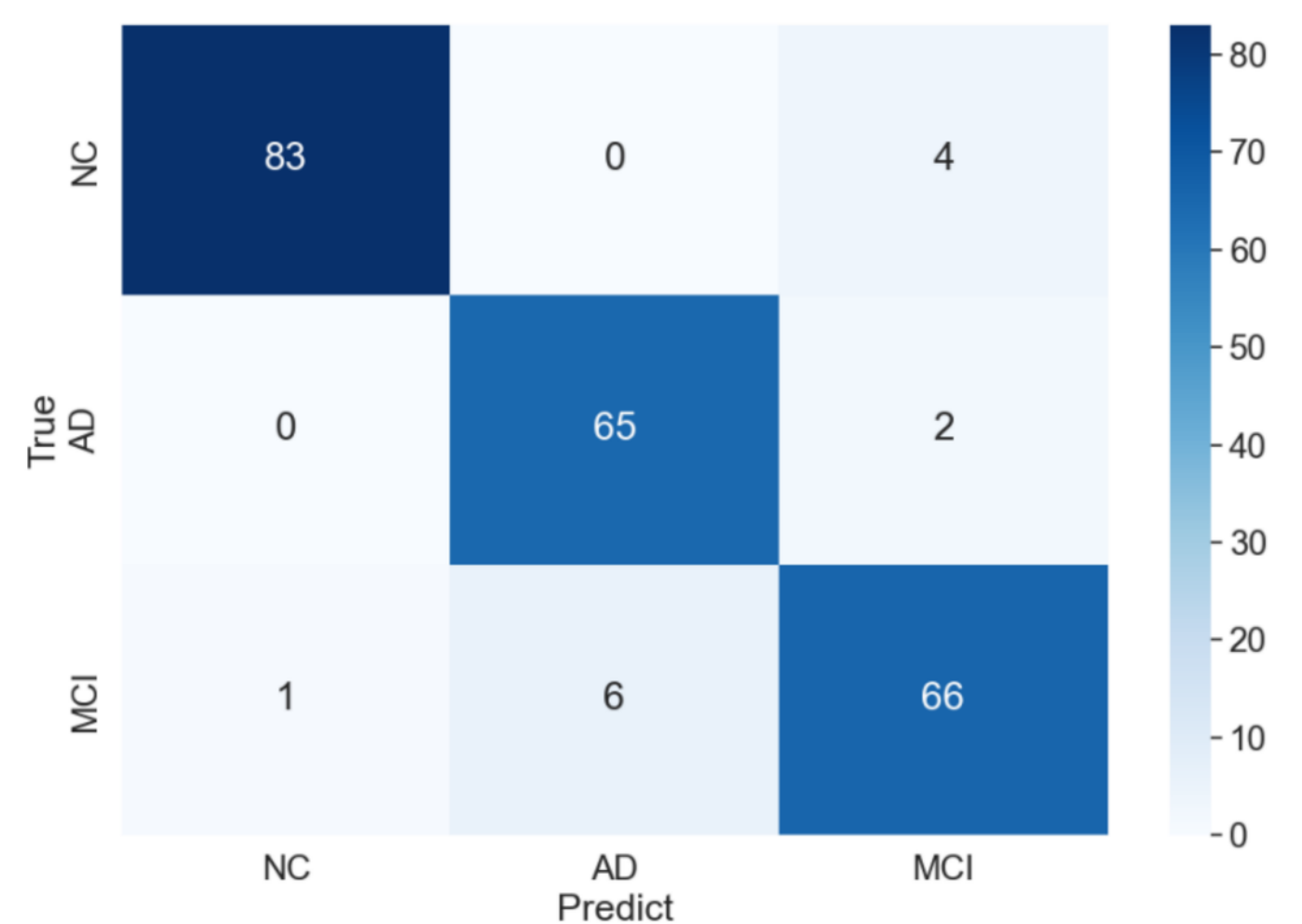



Table 1. Results of ablation experiments.

1      **Table 1.** Results of ablation experiments.

| Methods | Acc (%) | Sen (%) | Spe (%) | Pre (%) | F1 (%) |
|---------|---------|---------|---------|---------|--------|
| 3D | 76.9 | 76.6 | 88.7 | 77.2 | 76.7 |
| 2D | 89.4 | 90.2 | 94.8 | 90.8 | 89.8 |
| 2D+3D | 91.6 | 91.6 | 95.9 | 92.2 | 91.3 |
| SE+2D+3D | 89.8 | 89.6 | 95.1 | 89.6 | 89.5 |
| AMSF | **94.3** | **94.2** | **97.1** | **94.2** | **94.1** |

Table 2. The comparison results with other previous methods for AD vs. MCI vs. NC classification.

1    **Table 2.** The comparison results with other previous methods for AD vs. MCI vs. NC classification.

| Methods | Category with the number of samples | Acc(%) |
|---|---|---|
| Automatic CL | AD: 95,MCI: 207,NC: 104 | 87.2 |
| JMMP-LRR | AD: 24,MCI: 24,NC: 24 | 89.0 |
| DemNet | AD: 300,MCI: 300,NC: 300 | 91.9 |
| Linner SVM | AD: 34,MCI: 45,NC: 21 | 92.0 |
| AdaBoost | AD: 85,MCI: 193,NC: 111 | 92.7 |
| AMSF (ours) | AD: 301,MCI: 331,NC: 351 | **94.3** |