

## APPENDIX 1

### SEARCH STRATEGY FOR SYSTEMATIC REVIEW

#### RESEARCH QUESTION

What is the validity and accuracy of supervised and unsupervised segmentation methods of white matter hyperintensities on MR images?

#### STEP 1

Defined the review question as:

Patient population: Elderly population

Disease (Radiological manifestation): White matter hyperintensities

Index tests: Supervised and unsupervised segmentation methods

Reference test: Manual segmentation method

#### STEP 2

##### SUBJECT COMPONENTS

Following subject components identified from the research question: accuracy, supervised segmentation, unsupervised segmentation, white matter lesion, MRI

#### STEP 3

##### EXPANDING SUBJECT COMPONENTS

Identified related terms, alternative spelling, and synonyms with reference to each subject component. Key words related to each subject component are given in the table (1) below. Hadjarian et al.,( 1998) used the term model driven segmentation for supervised segmentation and data driven segmentation for unsupervised segmentation in the article on fusion of supervised and unsupervised methods for abnormal regions<sup>5</sup>. Therefore, these terms are included in the search strategy. Various pattern recognition methods are considered as supervised segmentation techniques. Supervised segmentation methods are divided into two categories<sup>4</sup>. They are: pattern recognition and algebraic methods. Pattern recognition techniques are further classified into para metric, non-parametric, and artificial neural networks<sup>4</sup>. All these terms related to supervised segmentations are used in the search strategy. Unsupervised segmentation methods are also known as clustering<sup>4</sup>. Examples of unsupervised segmentation methods are k – means and fuzzy C means(FCM)<sup>4</sup>. The keywords related to the unsupervised segmentation methods are also included in the search strategy.

What is the validity and accuracy of supervised and unsupervised segmentation methods of white matter hyperintensities on MR images?				
Accuracy	Supervised segmentation	Unsupervised segmentation	White matter lesion	Magnetic resonance imaging
1.Validity	1.Supervised image segmentation	1.Unsupervised image segmentation	1.White matter hyperintensities	1.MRI
2.Sensitivity			2.WMH	2.MR images
3.Specificity	2.Supervised method	2.Data driven segmentation	3.WMHs	3.Quantitative MRI
4.ROC			4.WML	4. T2 hyperintensities
5.ROC curve	3.Machine learning based segmentation	3.Fully automated segmentation	5.WMLs	
6.Dice coefficient	4.Deep learning	4.Computer aided detection	6.Periventricular white matter lesions	
7.Similarity index	5.Model driven segmentation	5.Automatic segmentation	7.PWMLs	
	6.Pattern recognition	6. Clustering	8.Deep white matter lesions	
	7. Artificial neural networks	7.classification	9.DWMLs	
	8.Parametric	8. FCM	10.White matter lesion size	
	9.Non-parametric	9.Fuzzy C means	11.Leukoaraiosis	
			12.Aging	
			13.White matter lesion volume	
			14.White matter disease	

**Table 1. Related terms of subject components**

Literature search conducted from 2015 until the present day on following databases: 1)Medline  
2) Science direct 3) Web of Science 4) IEEE Xplore

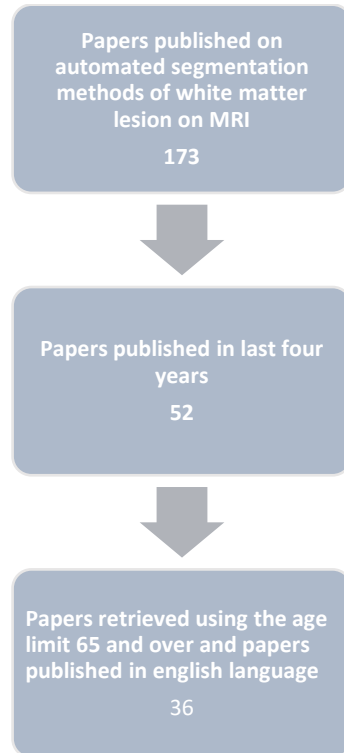
## SEARCH STRATEGY USED IN MEDLINE DATABASE

Ovid gateway is used to access the articles published on machine learning based segmentation methods of WMLs on MRI from 2015 to June 2019. Selected the resource Ovid Medline(R) and In-process & other non-indexed citations. Indexed term Leukoaraiosis is used to retrieve all the articles published about white matter lesions. Number of papers published in relation the review question is 173. Search results are narrowed down by selecting the articles published from 2015. The results are further narrowed down by restricting the age limits of the subjects to 65 and over.

- 1.expLeukoaraiosis
- 2.white matter hyperintensities\*.mp.
3. WMHs.tw.
4. white matter lesion\*.mp.
5. WMLs.tw.
6. Periventricular white matter lesions.tw.
- 7.PWMLs.tw.
8. Deep white matter lesions.tw.
9. DWMLs.tw.
10. white matter lesion size.tw.
11. aging.tw.
12. white matter lesion volume.tw.
13. White matter disease.tw.
14. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13
15. "Neural Networks (Computer)"/ or Pattern Recognition, Automated/ or Algorithms/ or Machine Learning/ or Image Processing, Computer-Assisted/ or Artificial Intelligence/
16. Supervised image segmentation.tw.
17. supervised method.tw.
18. Model driven segmentation.tw.
19. Artificial neural networks.tw.
20. Parametric segmentation.tw.
21. Non-parametric.tw.
22. Unsupervised image segmentation.tw.
23. machine learning based segmentation.tw.
24. data driven segmentation.mp.
25. fully automated segmentation.tw.
26. computer aided detection.tw.
27. automatic segmentation.tw.
28. clustering.tw.

29. classification.tw.
30. FCM.tw.
31. Fuzzy C means.tw.
32. deep learning.tw.
- 33.machine learning.tw.
- 34.deep learning based segmentation.tw
35. 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34
36. \*Magnetic Resonance Imaging/
37. MRI.tw.
38. MR images.tw.
39. Quantitative MRI.tw.
40. T2 hyperintensities.mp.
41. FLAIR hyperintensities.tw
- 42.36 or 37 or 38 or 39 or 40 or 41
43. validity.tw.
44. accuracy.tw.
45. similarity index.tw.
46. dice coefficient.tw.
47. sensitivity.tw.
48. specificity.tw.
49. ROC curve.tw.
50. ROC.tw.
51. 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50
51. 14 and 35 and 42 and 51
52. limit 52 to last 4 years
53. limit 53 to humans
54. limit 54 to "middle aged (45 plus years)"
- 55.limit 55 to English language

### Breakdown of Medline search results



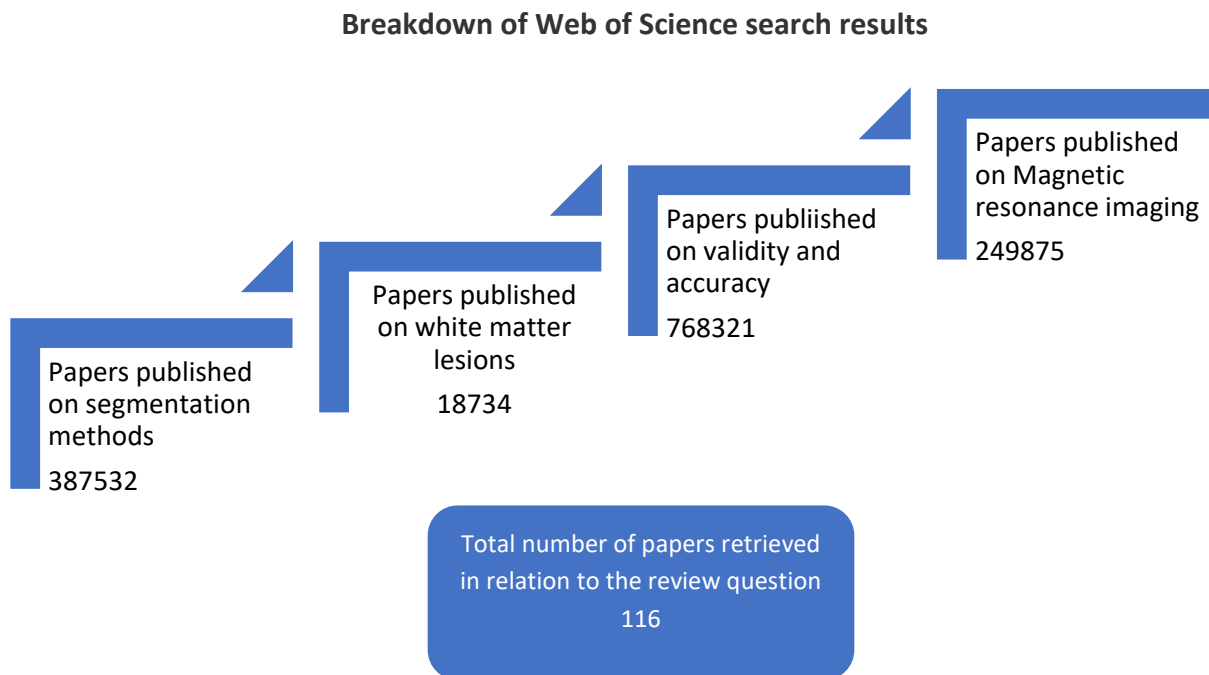
### SEARCH STRATEGY USED IN WEB OF SCIENCE DATABASE

Selected all databases on Web of science. Advanced search has been used to retrieve all the journal articles published on machine learning based segmentation methods of WMLs on MRI. Timespan is customised. Selected the time span from 2015 to 2019. Selected English language for the section 'search language to use'. Field tags TS and TI are used where TS stands for topic and TI stands for title. Boolean operators AND and OR are also used. Subject components identified above are used in the search strategy.

#### Search results

1. TI=supervised segmentation
2. TI=(unsupervised segmentation OR clustering OR classification )
3. TI=(FCM OR fuzzy C means OR computer aided detection OR fully automated segmentation)
4. TS=( pattern recognition OR fully automated segmentation OR supervised segmentation OR unsupervised segmentation OR model driven segmentation OR data driven segmentation OR artificial neural networks OR parametric OR non-parametric segmentation)

5. #4 OR #3 OR #2 OR #1
6. TS=(leukoaraiosis)
7. TS= (white matter hyperintensities\* OR WMHs OR white matter lesions OR WMLs OR periventricular white matter lesions\* OR deep white matter lesions\* OR white matter lesion volume OR White matter disease)
- 8.#7 OR #6
9. TS=( magnetic resonance imaging OR MRI OR MR images OR quantitative MRI)
10. TI=(T2 hyperintensities)
11. TI = (FLAIR hyperintensities)
12. #11 OR #10 OR #9
13. TI=( validity OR accuracy)
14. TS=( dice coefficient OR similarity index OR sensitivity OR specificity OR ROC curve)
15. #14 OR #13
16. #15 AND #12 AND #8 AND #5



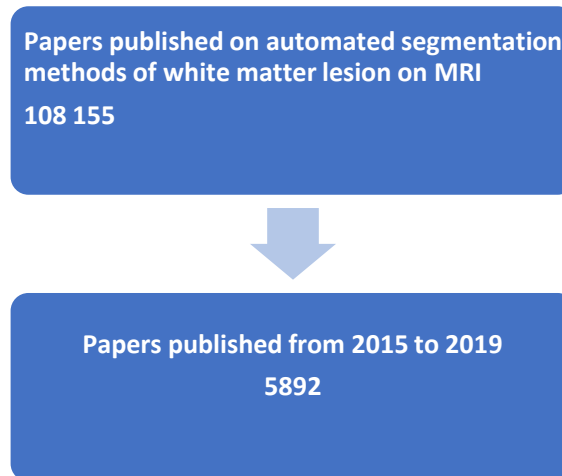
### SEARCH STRATEGY USED IN SCIENCE DIRECT

Advanced search is used to find as many articles published on supervised and unsupervised segmentation methods of WMLs on MRI. Selected research articles in the article types section. The advanced search does not allow to use Boolean operators more than 8 times in a single search. Hence, most relevant key words are used to retrieve papers published in relation to the review question.

Search terms used: (white matter hyperintensities OR white matter lesions OR T2 hyperintensities) AND (fully automated segmentation OR supervised segmentation OR unsupervised segmentation OR computer aided detection) AND (magnetic resonance imaging)

Number of papers published in relation to the research question is 5892.

#### Breakdown of Science Direct search results

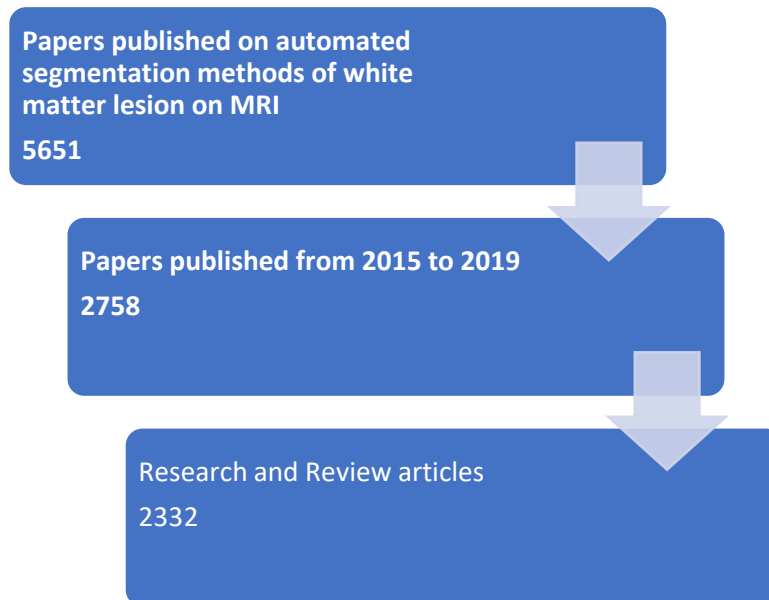


#### Refined search strategy for Science Direct

Number of papers retrieved using the search strategy above is 5892. The search results contained mostly irrelevant articles. The search strategy explained above is too broad and the search results retrieved all the articles containing at least one of the key words. Hence, the number of articles related to the research question increased significantly. Since, this is not a manageable number the search strategy is further refined. Developed a focused search strategy using the basic search option instead of advanced search.

Search terms used: (white matter hyperintensities OR white matter lesion) AND segmentation

### Breakdown of Science Direct search results after refining the search strategy



### SEARCH STRATEGY USED IN IEEE XPLORE

Selected advanced search to retrieve articles related to machine learning based segmentation techniques of WMLs on MR images. Radio button metadata only is selected in advanced search section. Publication year is customised for the period 2015 to 2019.

Key words: (white matter lesions OR white matter hyperintensities)

Combining the key word segmentation with above stated key words retrieved only 24 search results. Applying a focused search strategy may miss important articles published on segmentation methods of WMLs. Hence, the key words related to segmentation are not included in the search strategy used for IEEE Xplore.

**Search Results:** Number of articles retrieved from IEEE xplere is 64.

Mendeley reference manager is used to organise the articles retrieved from online databases. Duplicates have been removed and merged all the articles.



### Breakdown of the search results

Medline 36	Web of Science 116	Science Direct 2332	IEEE Xplore 64
	Duplicates 104	Papers retrieved after removing duplicates 2327	