An Image Patch Row-Column Ranking Method Using the Feature Accumulation Matrix to Explain Decisions of a Convolutional Neural Network

Luna M. Zhang
Department of Computer Science
Stony Brook University
Stony Brook, NY 11794

I. INTRODUCTION

- To sufficiently interpret decisions of a Convolutional Neural Network (CNN), we need to additionally use 1D image information associated with image patch rows and image patch columns. For example, a medical doctor may need to know about which image patch rows and image patch columns are most important for diagnosis.
- Since feature selection (FS) is useful in not only improving the model performance but also in interpreting a deep neural network, we develop an efficient and accurate CNN model by adding a novel layer called the "FS Layer."
- We create an image patch row-column ranking method to rank the top image patch rows and top image patch columns to use 1D image information for explaining decisions of the CNN model with the FS layer.

II. An Efficient CNN with FS and Row-Column-Ranking

Definition 1: Let the "feature accumulation matrix" A have elements called "feature accumulators" a_{ij} for $i = 0, 1, \ldots, H-1$ and $j = 0, 1, \ldots, W-1$, where $a_{ij} = \sum_{s=1}^{N} b_{ij}^{s}$ where b_{ij}^{s} is an element of the feature binary matrix B^s, and N is the number of feature maps.

Algorithm 1 Image Patch Row-Column Ranking

Input: A feature accumulation matrix A.

Output: Top N patch rows for $1 \leq N < H$ and top M patch columns $1 \leq M < W$.

1: Step 1: Calculate the patch row-wise feature accumulation number A_{row}^{i}: $A_{row}^{i} = \sum_{j=0}^{W-1} a_{ij}$ for $i = 0, 1, \ldots, H-1$.

2: Step 2: Sort $\{A_{row}^{0}, A_{row}^{1}, \ldots, A_{row}^{H-1}\}$ to get top N patch row numbers.

3: Step 3: Calculate the patch column-wise feature accumulation number A_{column}^{j}: $A_{column}^{j} = \sum_{i=0}^{H-1} a_{ij}$ for $j = 0, 1, \ldots, W-1$.

4: Step 4: Sort $\{A_{column}^{0}, A_{column}^{1}, \ldots, A_{column}^{W-1}\}$ to get top M patch column numbers.

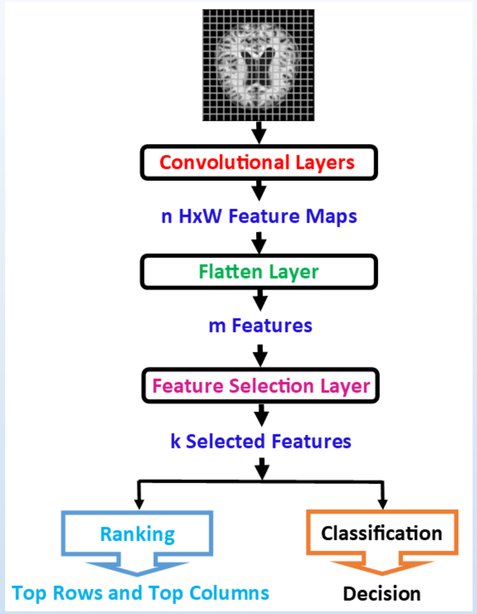

Fig. 1. The Features-based Row-Column-Ranking Algorithm

III. SIMULATION RESULTS

The Alzheimer's MRI preprocessed dataset with 6,400 128x128 images [1] are used for 4-class classification performance analysis. The fine-tuned ResNet50-FS model using 800 selected features with test accuracy of 0.9891 is better than the conventional ResNet50 using all 16,384 features with test accuracy of 0.9642.

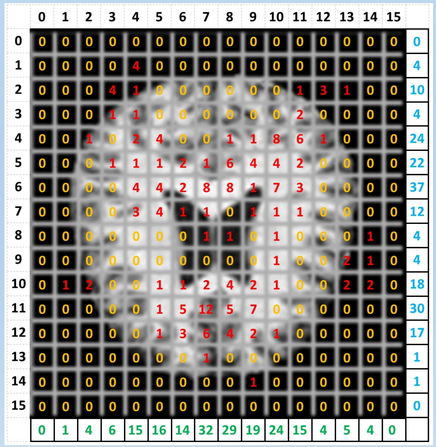

Fig. 2. A feature accumulation matrix (188 common features)

Table 1 shows that patch rows 6 and 11 have 6 brain areas among the 10 brain areas associated with AD and 9 brain areas among the 43 brain areas that may be associated with AD [2]. Table 1 also shows that patch columns 7 and 8 have 5 brain areas among the 10 brain areas associated with AD and 6 brain areas among the 43 brain areas that may be associated with AD. The top patch rows and the top patch columns are not associated with any brain areas that may not be associated with AD. A doctor may use 1D information in the top patch rows and top patch columns to partially interpret a diagnosis.

Row 6	Row 11	Column 7	Column 8
front to occi	front to occi	front to occi	front to occi
temporal to parietal	33 (ACC)	33 (ACC)	33 (ACC)
PFcm (IPL)	p24c	p24ab	p24ab
TPJ (STG/SMG)	p24ab	p24c	p24c
TE 2.2 (STG)	p32	p32	p32
PF (IPL)	45 (IFG)	Fp1	Fp1
PFm (IPL)	frontal-I	Fp2	Fp2
PFop (IPL)	IFS1 (IFS)	hOc1	hOc1
		hOc2	hOc2
		hOc3d	hOc3d
		frontal-I	frontal-I

Table 1. Brain Areas Related to Top Rows and Columns ("front to occi" = "frontal to occipital", red: associated with AD, blue: may be associated with AD).

IV. CONCLUSIONS & FUTURE WORKS

- The top-ranked patch rows 6 and 11 and the top-ranked patch columns 7 and 8 have the most top-ranked common features, 6 brain areas associated with AD, and 15 brain areas that may be associated with AD.
- The new patch row-column ranking method can generate useful 1D image information in the top rows and top columns to interpret decisions of a CNN.
- Heatmaps will be used to rank rows and columns. Image patches will be ranked to get useful 2D image information. The new relationship among AD, top patch rows, top patch columns, and top patches will be analyzed.
- It is useful to use hybrid 2D-1D information in important brain areas associated with the top patches, top patch rows, and top patch columns to make more robust and explainable decisions.

V. REFERENCES

[1] Alzheimer MRI Preprocessed Dataset, https://www.kaggle.com/datasets/sachinkumar413/\\alzheimer-mri-dataset, 2024.
[2] ebrains (Human Brain Atlas), https://www.ebrains.eu/tools/human-brain-atlas, 2024.