

## A DEEP LEARNING FRAMEWORK FOR BRAIN EXTRACTION IN HUMANS AND ANIMALS WITH TRAUMATIC BRAIN INJURY

Snehashis Roy<sup>†\*</sup> Andrew Knutsen<sup>†§</sup> Alexandru Korotcov<sup>†§</sup> Asamoah Bosomtwi<sup>†§</sup>  
Bernard Dardzinski<sup>§</sup> John A. Butman<sup>‡</sup> Dzung L. Pham<sup>†</sup>

<sup>†</sup> Center for Neuroscience and Regenerative Medicine, Henry Jackson Foundation

<sup>§</sup>Department of Radiology and Radiological Sciences,  
Uniformed Services University of the Health Sciences

<sup>‡</sup> Radiology and Imaging Sciences, Clinical Center, National Institute of Health

### ABSTRACT

Automatic brain extraction or skull stripping from magnetic resonance images (MRI) is an important pre-processing step in many image processing pipelines. Most skull stripping methods are optimized for normal brains and applicable to single  $T_1$ -w MR images. However, other contrasts, such as  $T_2$ , can provide complementary information about the boundary. This is especially true in the presence of traumatic brain injury (TBI) and other diseases, where lesions can confound boundary definitions. In this paper, we propose a deep learning based framework to extract intracranial tissues from multi-contrast MR images in the presence of TBI. Our approach is based on state-of-the-art convolutional neural network architecture to learn a transformation from multi-contrast atlas MR images to their stripping masks without using any deformable registration. An advantage of our framework is that it can be applied to different species. We applied our approach to 19 human patients with mild to severe TBI, as well as 16 normal mice images, and another 10 mice brains with TBI. We compared the approach with 3 separate state-of-the-art human and rodent brain extraction methods. Using only a few manually delineated atlases, we showed significant improvement in brain extraction accuracy in both healthy and pathological human and rodent images.

### 1. INTRODUCTION

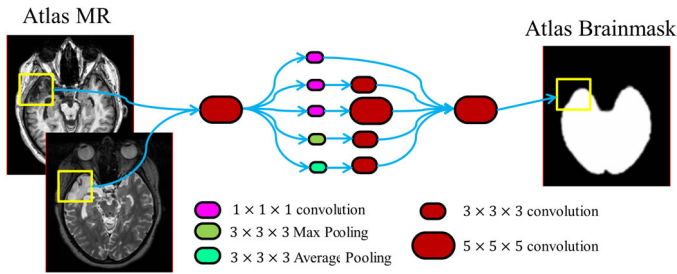
Automated brain extraction or skull stripping of whole head magnetic resonance (MR) images, which performs a binary classification of the image into brain and non-brain tissues, is often one of the earliest, yet most crucial, processing steps in brain imaging studies. Subsequent image processing, such as tissue segmentation, cortical reconstruction, or registration, are affected by the accuracy of the skull stripping. Many

skull stripping algorithms for healthy human MR brain images have been proposed in the literature. However, most are optimized for  $T_1$ -w, as these are typically high resolution and offer excellent soft tissue contrast in clinical studies. There are two primary categories of skull stripping methods, edge based and atlas based. Edge based stripping methods, such as BET [1], BSE [2], and AFNI `3dSkullStrip`, try to find a continuous edge between brain and skull based on intensity gradients. Other edge-based approaches include watershed [3] and graph cuts [4].

Although edge based skull stripping methods have been reasonably successful, atlas based methods were shown to be more accurate [5]. Usually an “atlas” consists of a  $T_1$ -w image and its manually drawn brain mask. Multi-atlas approaches typically deformably register multiple atlases to a subject and their corresponding brain masks are transformed to the subject space. The transformed atlas brain masks can be combined by majority voting [6, 7] or label fusion [8]. The final accuracy of these approaches depends largely on the accuracy of registrations. Another drawback of multi-atlas registration methods is that they require significant computational overhead due to a large number of deformable registrations. Alternately, patch based methods use only a few atlases and do not require accurate deformable registrations. Atlases are registered to the subject with only an affine [9, 10] or a coarse deformable registration [11]. Then the corresponding atlas brain masks are combined and iteratively refined with deformable models [9], intensity based segmentation [12], or patch-based label fusion [11, 10].

Much of the skull stripping literature has focused on human imaging, but skull stripping in small animal models remains an important and challenging problem, since most human brain skull stripping algorithms are not directly applicable to small animals. A modified version of BET [1] is provided with AFNI `3dSkullStrip` for rodent brains with `-rat` option. Morphological filtering with intensity gradients based deformable surface reconstruction has been used for brain extraction [14]. An unsupervised artificial neural

\*This work was supported by the Department of Defense in the Center for Neuroscience and Regenerative Medicine and the Intramural Research Program of the National Institutes of Health. This work was also partially supported by grant from National MS Society RG-1507-05243.



**Fig. 1.** The proposed CNN architecture using modified Google Inception [13] block is shown. Patches from an atlas MPRAGE and  $T_2$ -w images are passed through the Inception block, with the output being a patch of the atlas brain mask membership.

networks based method, called pulse coupled neural networks (PCNN) [15, 16], use iterative filtering and morphological operations to obtain brain mask.

Skull stripping methods in both human and animal imaging have generally focused on healthy brains devoid of pathology, despite the fact that there is obviously a greater need to study diseased brains. Our group showed that the use of MRI multiple contrasts offers an important advantage when performing skull stripping in brains with lesions [11]. Recently, deep learning [17], or convolutional neural networks (CNN), have been successfully used in many MR image processing applications. A CNN based stripping method [18] was proposed for brain extraction from images with tumors. In this paper, we propose a CNN based skull stripping framework to extract the intracranial volume from whole head MR images. We show that using the same framework on both human and mouse brains with TBI, significant improvement in brain extraction can be achieved over other state-of-the-art methods. Compared to previous methods, our approach has two key advantages. First, because of the design of the CNN architecture, only a few (2 – 4) atlases with manually delineated brainmasks are required. Second, accurate deformable registration is not needed, which reduces computational overhead and increases robustness to the presence of pathology.

## 2. DATASET

The proposed skull stripping method was evaluated on three sets of multi-contrast images. The first set consists of 19 patients with mild to severe TBI, having  $T_1$ -w MPRAGE (Siemens 3T,  $T_R = 2.53$  s,  $T_E = 3.03$  ms,  $T_I = 1.1$  s, flip angle  $7^\circ$ , resolution  $1 \times 1 \times 1$  mm<sup>3</sup>) and  $T_2$ -w fast spin echo images ( $T_R = 3.2$  s,  $T_E = 409$  ms, flip angle  $120^\circ$ , resolution  $0.98 \times 0.98 \times 1$  mm<sup>3</sup>). For training and evaluation, the  $T_2$ -w images were rigidly registered [19] to the MPRAGE, and binary brain masks were drawn on the registered  $T_2$ -w images. The second set consists of 16 normal mice and the third set contains 10 mice used in a closed head model of repetitive TBI (rTBI). Each mouse underwent 3 injuries using a controlled cortical impact de-

vice to study the physiological and behavioral impact. The mouse dataset consists of multi-echo 3D RARE scans with  $T_R/T_E = 4000/(10/30/50/70/90/110)$  ms, resolution of  $0.125 \times 0.125 \times 0.75$  mm<sup>3</sup> for normal and  $0.1 \times 0.1 \times 0.6$  mm<sup>3</sup> for rTBI. We used two echoes at  $T_E = 10$  and  $T_E = 70$  ms for our experiments. Brain masks were delineated on the first echo ( $T_E = 10$  ms). All human images were corrected once for intensity inhomogeneity with N4 [20], while the mice images, having more severe inhomogeneity, were corrected twice. The increased inhomogeneity in mice images is due to the use of surface coils for animal experiments.

## 3. METHOD

The proposed CNN based framework is shown in Fig. 1. The network learns an intensity transformation from MR images to the brainmask using multiple atlases. An atlas consists of multiple contrasts ( $T_1$ -w and  $T_2$ -w in the case of human subjects) and its manually drawn brainmask. We extract  $p_1 \times p_2 \times p_3$  patch-pairs from the two atlas MR images. For every atlas MR patch-pair, the network predicts a *membership* of the corresponding brain mask patch. The training data consists of MR patch-pairs as well as the target brain mask membership patch. The membership of the brainmask is generated simply by applying a  $3 \times 3 \times 3$  Gaussian filter to the binary segmentation. As we are interested in the boundary between subarachnoid space and the meninges and skull, a narrow band region of width 5 voxels is first defined on the brain mask boundary. Then training patches from MR images and brainmask membership are extracted in such a way that the center voxel of a patch lies within that narrow region of interest. This ensures that all the training patches contain relevant information about the brain boundary.

We employed a recent state-of-the-art CNN architecture called Inception, proposed by Google [13]. Traditional CNN models, such as AlexNet [21], employs serial combinations of convolution, pooling, and downsampling/upsampling layers. The number of free parameters can increase exponentially with the number of layers. This can cause optimization instability when the number of training data becomes fewer than the number of parameters. The Inception block, shown in Fig. 1 tries to resolve this issue by parallelizing the convolution and pooling layers as well as employing downsampling by introducing a  $1 \times 1 \times 1$  (denoted  $1^3$ ) convolution. This operation essentially downsamples the number of filters from the previous layer. The originally proposed Inception block [13] contains parallel pathways to  $3^3$  and  $5^3$  convolution filters as well as maximum pooling, preceded by  $1^3$  convolutions. We modified the Inception architecture by adding a concatenation of  $3^3$  average pooling layers and  $3^3$  convolutions, which acts as a downsampling layer without actually decreasing the size of patch. Therefore the size of the patch before and after every layer is identical to the input MR and output brainmask patch size. The proposed network is fully convolutional, i.e., no fully connected layers are included, thereby reducing the number of free parameters.

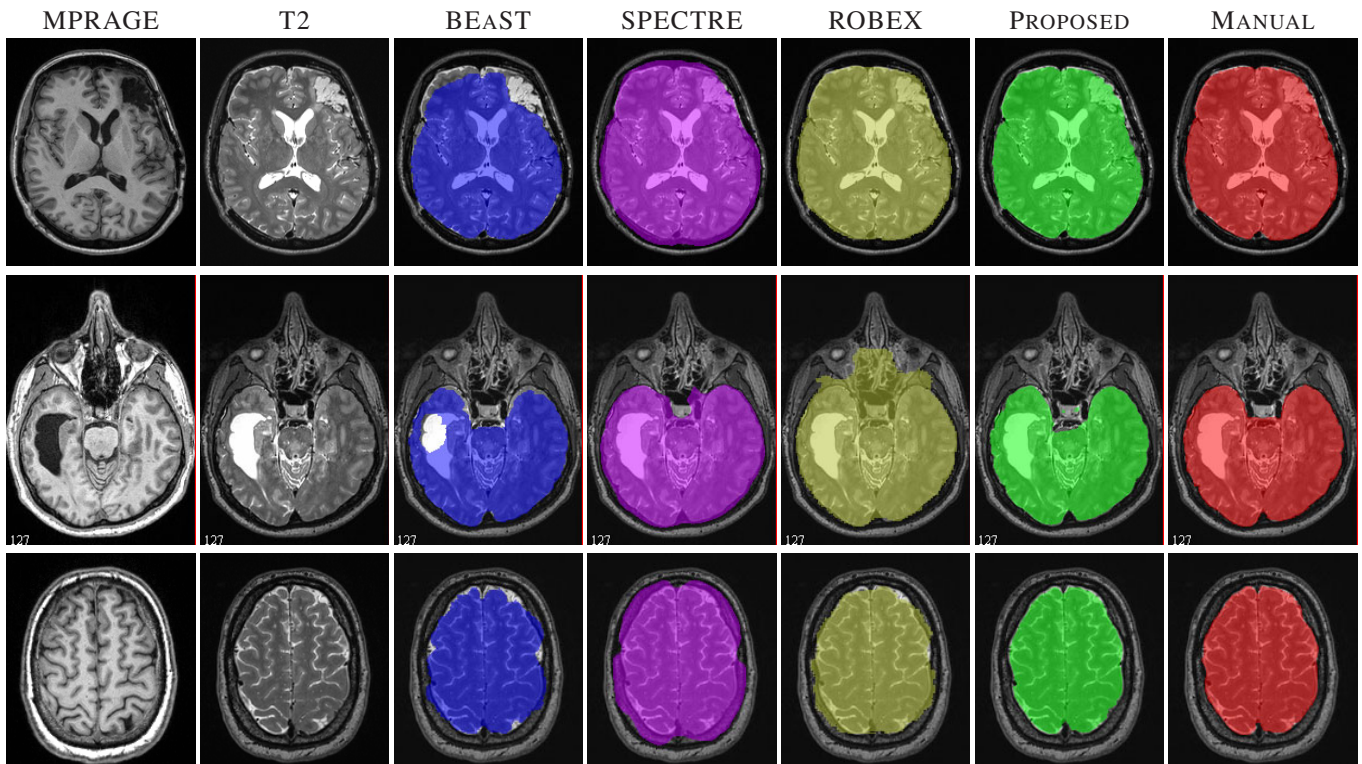


Fig. 2. Example brainmasks of 3 subjects are shown for 4 competing methods, BEaST [10], ROBEX [9], SPECTRE [12], and the proposed CNN based algorithm.

The model is generalizable to both 2D and 3D patches. The number of filters in the first  $5^3$  convolution is 128, while the number of filters in the Inception block is kept the same as proposed in the original paper [13]. We added the average pooling layer followed by 32 filters of size  $3^3$ . The Inception block is followed by one  $5^3$  filter to keep the input and output patch size consistent. The filter parameters are learned by stochastic gradient descent via Adam [22]. For the human TBI brains, we used 3D  $19 \times 19 \times 19$  patches since the images were isotropic. For mice brains, we used 2D  $41 \times 41$  patches for both normal and rTBI mice, since the slice thickness is 6 times the in-plane resolution. The learnt 2D models are applied slice by slice on the mice images.

#### 4. RESULTS

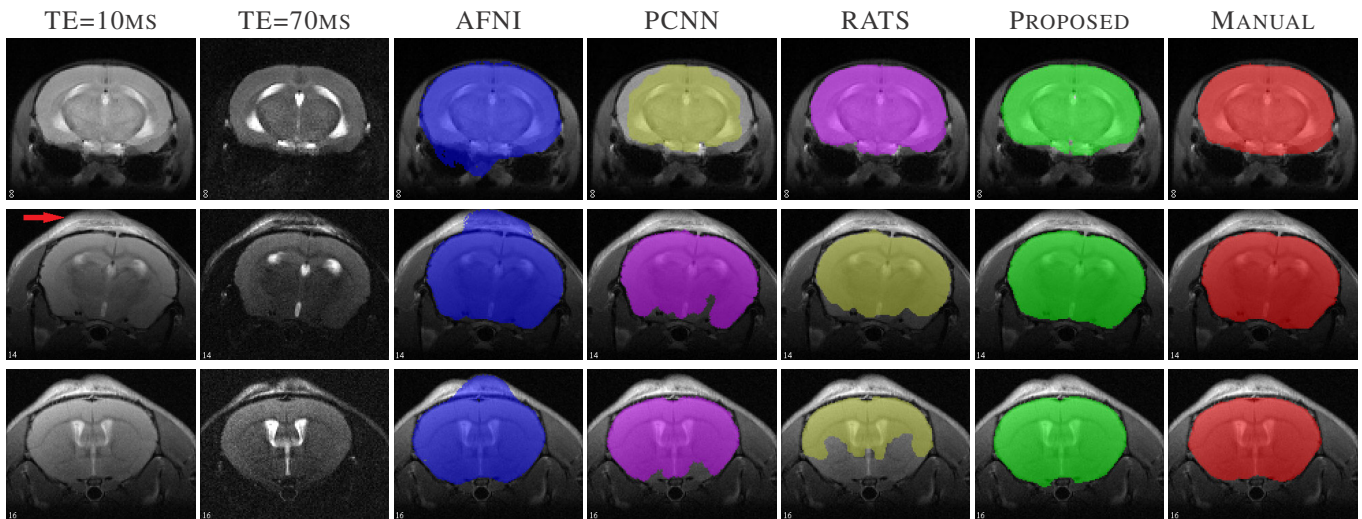
For the human TBI dataset, we arbitrarily chose 3 subjects as atlases, and trained the network separately for each of the atlases to generate 3 trained models. Then for every subject, each of the trained models are applied to generate 3 memberships. The memberships are averaged and thresholded at 0.5 to obtain a binary segmentation. We compared with 3 other state-of-the-art atlas based stripping methods, BEaST [10], ROBEX [9], and SPECTRE [12]. All of them use  $T_1$ -w images to estimate the mask. Examples of 3 subjects are shown in Fig. 2. The top two rows show subjects with a hemorrhage and enlarged ventricles, where BEaST consistently underestimates the mask, and SPECTRE overestimates it. The

Table 1. Quantitative comparison with respect to Dice, Jaccard, positive predictive value (PPV), and volume difference (VD) between 4 competing methods on the three dataset, human TBI, normal mice, and rTBI mice. Bold indicates significant improvement ( $p < 0.05$ ) over all the other methods.

Dataset	Method	Metrics			
		Dice	Jacc	PPV	VD
Human TBI	BEaST	0.8848	0.8036	0.9348	0.1284
	SPECTRE	0.9002	0.8295	0.9054	0.0745
	ROBEX	0.8866	0.8045	0.8380	0.1394
	Proposed	<b>0.9719</b>	<b>0.9454</b>	<b>0.9963</b>	<b>0.0477</b>
Normal Mice	AFNI	0.9067	0.8488	0.8731	0.1000
	PCNN	0.8573	0.7606	0.7739	0.2683
	RATS	0.8459	0.7406	0.8905	0.0770
	Proposed	<b>0.9486</b>	<b>0.9022</b>	<b>0.9436</b>	<b>0.0196</b>
rTBI Mice	AFNI	0.9217	0.8450	0.9090	0.0420
	PCNN	0.9160	0.8486	0.9503	0.0719
	RATS	0.8405	0.7309	0.9410	0.1830
	Proposed	<b>0.9543</b>	<b>0.9128</b>	0.9612	0.0329

third row shows a subject where the dura and marrow have a similar intensity as GM, which is why both SPECTRE and ROBEX include it inside brainmask. Our proposed method,





**Fig. 3.** Example brain masks of 3 mouse images are shown for 4 competing methods: AFNI `3dSkullStrip`, PCNN [16], RATS [14], and the proposed method. Top row shows a normal mouse brain, while the bottom two rows have rTBI.

using multi-contrast images, shows better stripping masks in all 3 subjects.

For both mice datasets, the binary brainmasks were generated as before. We compared to AFNI `3dSkullStrip` with `-rat` option enabled, PCNN [16], and RATS [14]. We used 4 and 2 atlases for normal and rTBI mice, and generated brainmasks on 12 normal and 8 rTBI mice, respectively. Fig. 3 shows examples of a normal (top row) and two rTBI mice (middle and bottom row). AFNI and RATS perform similar to the proposed method on the normal mouse, while PCNN shows some underestimation. Note that for the rTBI cases, the point of injury is near the frontal lobe, where the tissue is swollen (red arrow). AFNI includes part of the frontal lobe, while PCNN and RATS shows significant underestimation. The proposed method provides more accurate segmentations in each case.

For quantitative comparison, we used Dice, Jaccard, positive predictive value (PPV), and volume difference (VD). Table 1 shows the mean values of the metrics on all of the three datasets. For human TBI data, the proposed method showed significant improvement ( $p < 0.05$ ) over all the competing methods for all 4 metrics. For the normal mouse dataset, the proposed method achieved the highest Dice, Jaccard, and PPV compared to the other three methods ( $p < 0.05$ ). Although the volume difference is lowest for the proposed method, it is not significant w.r.t. RATS. Similar performance was observed in the rTBI data, where our method achieved significantly higher Dice and Jaccard ( $p < 0.05$ ), while PPV for PCNN and VD of AFNI are statistically not different from our method. Note that our method had the highest Dice, Jaccard, PPV, and lowest VD among all the methods.

## 5. DISCUSSION

We have proposed a CNN based brain extraction framework for skull stripping based on multi-contrast MR images in the presence of TBI in both humans and small animals. Experiments on both human and mouse MR images show significant improvement in skull stripping accuracy over 3 state-of-the-art methods. The primary advantage of our method is twofold. First, we showed significant improvement with just 3 atlases, whereas BEaST uses 10. Similarly, only two atlases were used in rTBI data. Second, we did not need any deformable registration between subject and atlas. While the training with one atlas takes about 10 hours, prediction of a new brainmask takes less than a minute on an Nvidia Titan X GPU.

On average, all methods performed better on rTBI mice images compared to the normal ones. We hypothesize that this is partially due to the fact that inhomogeneity was not completely removed from all the normal mice images. Significant signal loss was observed near the posterior regions, causing all methods to either overestimate or underestimate in that region. Also PCNN and RATS need initial thresholds for intensities and volume, which may not be optimal for all datasets, as MR intensities are not standardized. Further exploration is needed to fully identify the performance difference. Our future work include determination of the optimal number of atlases and patch size. We plan to release this software freely for public use.

## 6. REFERENCES

- [1] M. Jenkinson and S.M. Smith, "A global optimisation method for robust affine registration of brain images," *Med. Image Anal.*, vol. 5, no. 2, pp. 143–156, 2001.

- [2] D. Shattuck, S. Sandor-Leahy, K. Schaper, D. Rottenberg, and R. Leahy, "Magnetic resonance image tissue classification using a partial volume model," *NeuroImage*, vol. 13, no. 5, pp. 856–876, 2001.
- [3] F. Segonne, A. Dale, E. Busa, M. Glessner, D. Salat, H. K. Hahn, and B. Fischl, "A hybrid approach to the skull stripping problem in MRI," *NeuroImage*, vol. 22, no. 3, pp. 1060–1075, 2004.
- [4] S. A. Sadanathan, W. Zheng, M. W. L. Chee, and V. Zagorodnov, "Skull stripping using graph cuts," *NeuroImage*, vol. 49, no. 1, pp. 225–239, 2010.
- [5] R. Souza, O. Lucena, J. Garrafa, D. Gobbi, M. Saluzzi, S. Appenzeller, L. Rittner, R. Frayne, and R. Lotufo, "An open, multi-vendor, multi-field-strength brain MR dataset and analysis of publicly available skull stripping methods agreement," *NeuroImage*, pp. 1–13, 2017.
- [6] J. Doshi, G. Erus, Y. Ou, B. Gaonkar, and C. Davatzikos, "Multi-atlas skull-stripping," *Academic Radiology*, vol. 20, no. 12, pp. 1566–1576, 2013.
- [7] R. A. Heckemann, C. Ledig, K. R. Gray, P. Aljabar, D. Rueckert, J. V. Hajnal, and A. Hammers, "Brain extraction using label propagation and group agreement: Pincram," *PLoS One*, vol. 10, no. 7, pp. e0129211, 2015.
- [8] B. B. Avants, N. J. Tustison, G. Song, P. A. Cook, A. Klein, and J. C. Gee, "A reproducible evaluation of ANTs similarity metric performance in brain image registration," *NeuroImage*, vol. 54, no. 3, pp. 2033–2044, 2011.
- [9] J. E. Iglesias, C. Y. Liu, P. Thompson, and Z. Tu, "Robust brain extraction across datasets and comparison with publicly available methods," *IEEE Trans. Med. Imag.*, vol. 30, no. 9, pp. 1617–1634, 2011.
- [10] S. F. Eskildsen, P. Coupe, V. Fonov, J. V. Manjon, K. K. Leung, N. Guizard, S. N. Wassef, L. R. Ostergaard, D. L. Collins, and The Alzheimer's Disease Neuroimaging Initiative, "BEaST: Brain extraction based on nonlocal segmentation technique," *NeuroImage*, vol. 59, no. 3, pp. 2362–2373, 2012.
- [11] S. Roy, J. A. Butman, D. L. Pham, and Alzheimers Disease Neuroimaging Initiative, "Robust skull stripping using multiple MR image contrasts insensitive to pathology," *NeuroImage*, vol. 146, pp. 132–147, 2017.
- [12] A. Carass, J. Cuzzocreo, M. B. Wheeler, P. L. Bazin, S. M. Resnick, and J. L. Prince, "Simple paradigm for extra-cerebral tissue removal: Algorithm and analysis," *NeuroImage*, vol. 56, no. 4, pp. 1982–1992, 2011.
- [13] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Intl. Conf. on Comp. Vision. and Patt. Recog. (CVPR)*, 2015, pp. 1–9.
- [14] I. Oguz, H. Zhang, A. Rumble, and M. Sonka, "RATS: Rapid automatic tissue segmentation in rodent brain MRI," *J. of Neuroscience Methods*, vol. 221, pp. 175–182, 2014.
- [15] M. Murugavel and J. M. Sullivan, "Automatic cropping of MRI rat brain volumes using pulse coupled neural networks," *NeuroImage*, vol. 45, no. 3, pp. 845–854, 2009.
- [16] N. Chou, J. Wu, J. B. Bingren, A. Qiu, and K.-H. Chuang, "Robust automatic rodent brain extraction using 3-D pulse-coupled neural networks (PCNN)," *IEEE Trans. Med. Imag.*, vol. 20, no. 9, pp. 2554–2564, 2011.
- [17] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [18] J. Kleesiek, G. Urban, A. Hubert, D. Schwarz, K. Maier-Hein, M. Bendszus, and A. Biller, "Deep MRI brain extraction: A 3D convolutional neural network for skull stripping," *NeuroImage*, vol. 129, pp. 460–469, 2016.
- [19] B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee, "Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain," *Med. Image Anal.*, vol. 12, no. 1, pp. 26–41, 2008.
- [20] N. J. Tustison, B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich, and J. C. Gee, "N4ITK: improved N3 bias correction," *IEEE Trans. Med. Imag.*, vol. 29, no. 6, pp. 1310–1320, 2010.
- [21] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems (NIPS)*, pp. 1097–1105, 2012.
- [22] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Intl. Conf. on Learning Representations (ICLR)*, 2015.