

Automatic Assessment of ASPECTS Using Diffusion-Weighted Imaging in Acute Ischemic Stroke Using Recurrent Residual Convolutional Neural Network

Luu-Ngoc Do ^{1,†}, Byung Hyun Baek ^{1,2,†}, Seul Kee Kim ^{1,3}, Hyung-Jeong Yang ^{4,5,*}, Ilwoo Park ^{1,2,5,*} and Woong Yoon ^{1,2}

¹ Department of Radiology, Chonnam National University, Gwangju 61469, Korea; doluungoc@gmail.com (L.-N.D.); qorqod@hanmail.net (B.H.B.); kimsk.rad@gmail.com (S.K.K.); radyoon@jnu.ac.kr (W.Y.)

² Department of Radiology, Chonnam National University Hospital, Gwangju 61469, Korea

³ Department of Radiology, Chonnam National University Hwasun Hospital, Hwasun 58128, Korea

⁴ Department of Electronics and Computer Engineering, Chonnam National University, Gwangju 61186, Korea

⁵ Department of Artificial Intelligence Convergence, Chonnam National University, Gwangju 61186, Korea

* Correspondence: hjyang@chonnam.ac.kr (H.-J.Y.); ipark@jnu.ac.kr (I.P.); Tel.: +82-62-530-3436 (H.-J.Y.); +82-62-220-5744 (I.P.)

† These authors contributed equally to this work.

Received: 12 August 2020; Accepted: 07 October 2020; Published: 9 October 2020

Abstract: The early detection and rapid quantification of acute ischemic lesions play pivotal roles in stroke management. We developed a deep learning algorithm for the automatic binary classification of the Alberta Stroke Program Early Computed Tomographic Score (ASPECTS) using diffusion-weighted imaging (DWI) in acute stroke patients. Three hundred and ninety DWI datasets with acute anterior circulation stroke were included. A classifier algorithm utilizing a recurrent residual convolutional neural network (RRCNN) was developed for classification between low (1–6) and high (7–10) DWI-ASPECTS groups. The model performance was compared with a pre-trained VGG16, Inception V3, and a 3D convolutional neural network (3DCNN). The proposed RRCNN model demonstrated higher performance than the pre-trained models and 3DCNN with an accuracy of 87.3%, AUC of 0.941, and F1-score of 0.888 for classification between the low and high DWI-ASPECTS groups. These results suggest that the deep learning algorithm developed in this study can provide a rapid assessment of DWI-ASPECTS and may serve as an ancillary tool that can assist physicians in making urgent clinical decisions.

Keywords: deep learning; diffusion magnetic resonance imaging; stroke

1. Introduction

Acute ischemic stroke is a major cause of disability, and urgent decision making for proper treatment strategies is critical for the improved outcome of patients with stroke [1,2]. The evaluation of infarct volume and stroke onset time play important roles in the early treatment approach [2]; hence the early detection and rapid quantification of the acute ischemic lesion on brain imaging with computerized tomography (CT) or diffusion-weighted magnetic resonance imaging (DWI) have become important for the diagnosis and treatment of acute ischemic stroke.

The Alberta Stroke Program Early CT Score (ASPECTS) is an established 10-point semi-quantitative scoring system using brain CT and has been used for the rapid assessment of the extent

of early ischemic changes in patients with acute anterior circulation stroke. Although the ASPECTS system has been widely utilized to determine the eligibility criteria of mechanical thrombectomy [3,4], the lack of agreement and the variability of ASPECTS among even experienced clinicians have been the main source of its limitation [5]. The DWI-ASPECTS, which is based on an ASPECTS measurement using DWI instead of CT, has been suggested as an alternative and shown to provide a superior inter-rater agreement and output prediction compared to CT-ASPECTS [6,7].

Recently, deep learning has been shown to be a powerful tool in computer vision, voice recognition, and natural-language processing, and has gained widespread attention for application in medical research. Convolutional neural networks (CNNs), a class of deep learning methods that has emerged as one of the most effective tools for pattern recognition, has been applied to the analysis of medical imaging, such as disease classification, lesion detection, segmentation, and data processing [8–10]. Although several recent studies have utilized deep learning algorithms to apply to CT and magnetic resonance imaging (MRI) data from stroke patients [11,12], the automated assessment of the extent of acute cerebral ischemia is still a challenging endeavor.

Building on previous work in which a 3D convolutional neural network (3DCNN) was applied for the binary assessment of DWI-ASPECTS [13], we developed a deep learning algorithm utilizing a recurrent residual convolutional neural network (RRCNN) for the automatic binary classification of DWI-ASPECTS from patients with acute anterior circulation ischemic stroke. We showed that the rapid assessment of DWI-ASPECTS using our proposed algorithm may provide a useful tool for physicians for categorizing these patients, which is of importance for time-sensitive clinical decision making.

2. Materials and Methods

2.1. Subjects

This retrospective study was approved by the Institutional Review Board of Chonnam National University Hospital, which waived the requirement for obtaining written consent. The data collection and all methods were carried out in accordance with the relevant guidelines and regulations. A total of 319 DWI datasets were included from patients who presented with acute anterior circulation stroke due to large vessel occlusion within 6 h of symptom onset at a tertiary stroke center from December 2010 and January 2016. An additional DWI dataset consisting of a total of 71 patients from February 2016 to November 2016 were collected for an independent test set. All patients underwent MRI examination using a 1.5T MRI scanner (Signa HDxt; GE Healthcare, Milwaukee, Wisconsin). DWI sequences were obtained in the axial plane by using a single-shot, spine-echo echoplanar technique with the following parameters: repetition time of 9000 ms, echo time of 80 ms, slice thickness of 4 mm, intersection gap of 0 mm, field of view (FOV) of 260×260 mm, matrix size of 128×128 (approximately $2 \text{ mm} \times 2 \text{ mm}$ in-plane resolution) and b-values of 0 and 1000 s/mm^2 .

DWI-ASPECTS was retrospectively assessed by two neuroradiologists who were blinded to clinical information. The assessment of DWI-ASPECTS involves 10 distinct regions, subdividing the territory of the middle cerebral artery, in which the overall score is determined by deducting a score of 1 from the initial score of 10 for each affected region [14]. Approximately 95% of the DWI-ASPECTS assessment (370 out of 390 cases) were the same between the two neuroradiologists. The final conclusions for the disagreeing 20 cases were made by a consensus from the two neuroradiologists after a further inspection by them. Patients were classified into two groups according to their DWI-ASPECTS: group 1 consisted of patients with the DWI-ASPECTS of 1–6 ($n = 147$) and group 2 with the DWI-ASPECTS of 7–10 ($n = 243$). This binary classification was based on the previous finding, which demonstrated the distinct clinical outcome between the two groups and consequently the need for rapid determination of differential treatment options between these two groups [15,16]. Figure 1 shows an example of DWI images from the two groups.

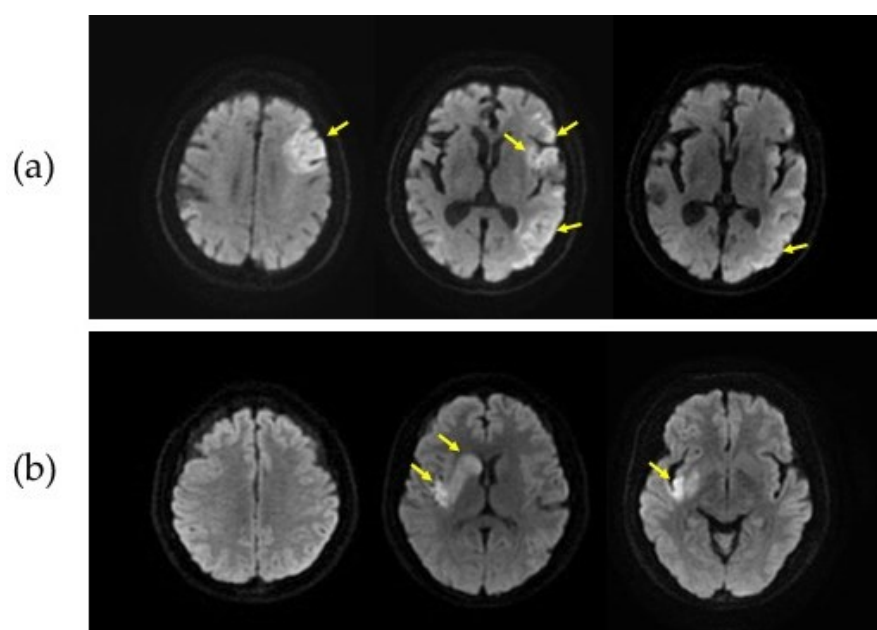


Figure 1. Representative diffusion-weighted magnetic resonance imaging (DWI) slices from two groups. DWI data from a patient with the Alberta Stroke Program Early Computed Tomographic Score (ASPECTS) of 5 (a) and the ASPECTS of 7 (b) are shown. The slices in the first, second and third column in each patient correspond to the supraganglionic, ganglionic, and infraganglionic levels, respectively. The yellow arrows indicate infarct lesions.

Eighty percent of the 319 DWI data points were randomly selected for the training dataset ($n = 244$) and the remaining 20% were kept for validation ($n = 75$). The additional dataset ($n = 71$) was used for the independent testing. The distribution of the ASPECTS among the training, validation, and independent test set is shown in Table 1.

Table 1. Distribution of the Alberta Stroke Program Early Computed Tomographic Score (ASPECTS) among the training, validation, and independent test set.

| | ASPECTS | | | | | | | | | | Group 1 (ASPECTS 1–6) | Group 2 (ASPECTS 7–10) | Total |
|------------|---------|---|---|----|----|----|----|----|----|----|--------------------------|---------------------------|-------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | | | |
| Training | 1 | 4 | 5 | 13 | 35 | 32 | 48 | 65 | 39 | 2 | 90 | 154 | 244 |
| Validation | 0 | 0 | 2 | 3 | 12 | 9 | 14 | 20 | 14 | 1 | 26 | 49 | 75 |
| Testing | 0 | 3 | 4 | 7 | 5 | 12 | 9 | 19 | 9 | 3 | 31 | 40 | 71 |

2.2. Slices Filtering

A slice filtering strategy was applied to bypass DWI slices that were non-informative. Each DWI dataset contained a sequence of approximately 40 imaging slices. Approximately 15% of cranial and 35% of caudal slices of each patient data, which were well outside the area of the middle cerebral arterial territory and uninformative for estimating ASPECTS, were removed. As a result, approximately 20 slices, with regions that included the middle cerebral arterial territory, remained.

2.3. DWI Preprocessing

After the slice filtering, the DWI datasets were preprocessed by brain cropping and a contrast stretching (Figure 2). The brain cropping was intended to remove the background portion of the images and to include only brain parenchyma. Based on the pixel intensity values from the image histogram, the top, bottom, left and right boundary pixels of the brain parenchyma were determined and utilized for selecting the boundary for cropping. The contrast of the cropped image was enhanced by the contrast stretching algorithm using the following formula:

$$P_{out} = (P_{in} - P_{in}^{min}) \left(\frac{P_{out}^{max} - P_{out}^{min}}{P_{in}^{max} - P_{in}^{min}} \right) + P_{out}^{min}$$

where P_{out} is the pixel value of the contrast-stretched image and P_{in} is the pixel value of the original image. P_{out}^{max} and P_{out}^{min} are 255 and 0, respectively. P_{in}^{max} and P_{in}^{min} are the 99th and 80th percentile of the histogram in the original image, respectively.

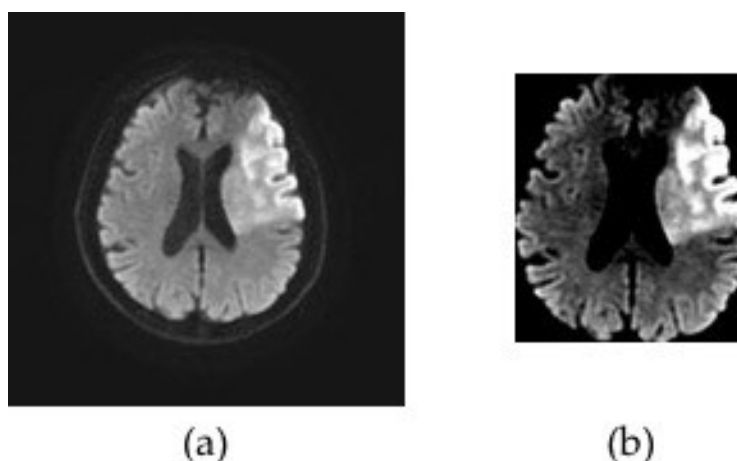


Figure 2. Representative diffusion-weighted magnetic resonance imaging (DWI) images illustrating the preprocessing step. The original DWI image (a) was cropped around brain parenchyma and a contrast stretching algorithm was applied to enhance the contrast of the DWI data (b).

2.4. Data Augmentation

Every 2D imaging slice of the DWI datasets was resized to 80×80 by a bicubic interpolation, rendering a final input DWI sequence with a $20 \times 80 \times 80$ resolution. The DWI datasets used for training the model were augmented by a horizontal flip, the addition of Gaussian noise, and a clockwise and counter-clockwise 15-degree rotation as shown in Figure 3. After the data augmentation, the training set consisted of 1220 DWI samples, corresponding to a total of 24,400 imaging slices.

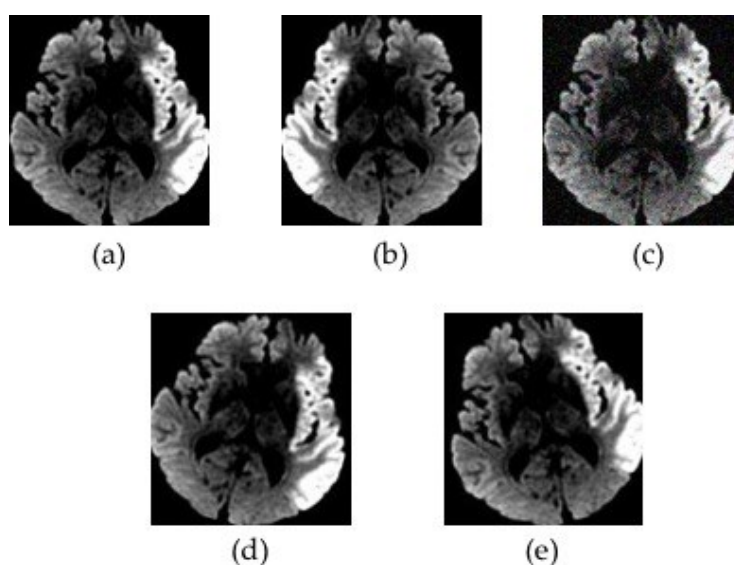


Figure 3. An illustration of data augmentation. The preprocessed diffusion-weighted magnetic resonance imaging data (a) underwent a horizontal flip (b), the addition of Gaussian noise (c), a clockwise 15-degree rotation (d), and a counter-clockwise 15-degree rotation (e).

2.5. Model Training

We used the RRCNN for model training, which considered the DWI slices as a sequence of images. It contained five convolution blocks for extracting 2D features and one recurrent block for extracting sequential features. The proposed RRCNN structure was adapted from the VGG16 [17] and ResNet [18] structure by adding skip connections in each convolution block as shown in Figure 4. Each block of convolution had two or three convolution layers with the kernel size of 3×3 or 7×7 and one max pooling layer. The number of feature maps in each convolution block was 32, 64, 128, 256, and 256, respectively. The recurrent block contained one long short term memory (LSTM) layer with 256 hidden nodes. The DWI datasets underwent a slice filtering, preprocessing, and augmentation step before training. We used an Adam optimizer with a learning rate of $1e-6$ and a batch size of 32. Our models were trained using Tensorflow-GPU with NVIDIA GTX 1080 Ti.

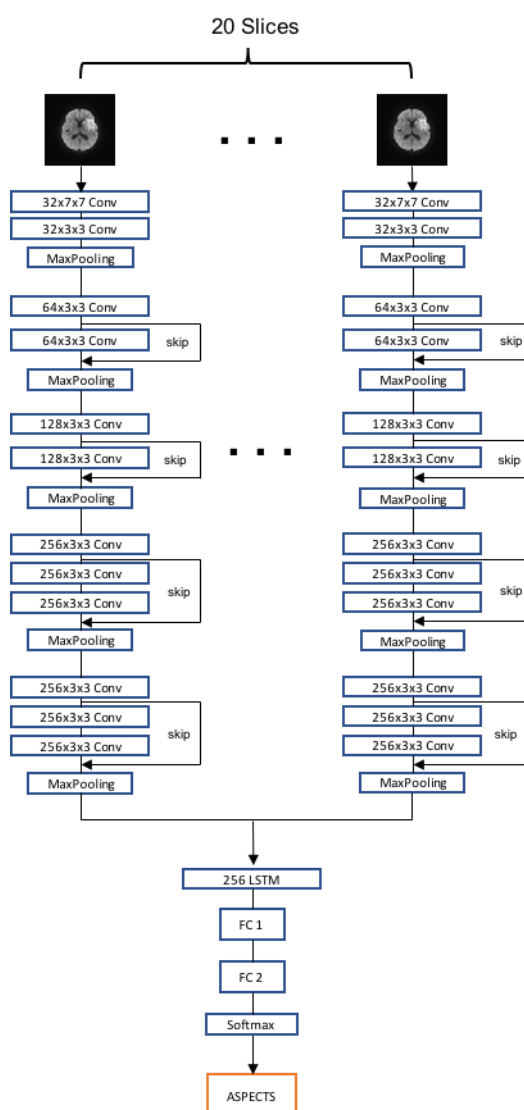


Figure 4. The flowchart of the recurrent residual convolutional neural network model developed in this study. The proposed model contained five convolution blocks followed by a recurrent block, fully connected layers, and a softmax classifier. LSTM, long short term memory; FC, fully connected layer; ASPECTS, Alberta Stroke Program Early Computed Tomographic Scoring.

2.6. Comparison of RRCNN with Pre-Trained Models and 3DCNN

We evaluated and compared the performance of the proposed RRCNN structure with those of VGG16 [17] and Inception V3 [19] models that were pre-trained using ImageNet data [20]. The pre-trained models were fine-tuned by adding one LSTM layer before the fully connected layers. During

training, the weights on the LSTM layer and the fully connected layers were updated, while the pre-trained CNN weights were frozen. We used an Adam optimizer with a learning rate of $1e-5$ and a batch size of 32. In addition, we trained the 3DCNN model [21], which considered the DWI slices as three-dimensional data, and compared its performance with that of the proposed RRCNN model. The learning rate and the batch size of 3DCNN was $1e-5$ and 32, respectively. Detailed information about the 3DCNN model has been previously described [13]. All deep learning models were trained and evaluated three times using the validation set in order to assess and verify the performance of their training and the mean performances were reported. Sensitivity (Group 1 considered as a positive condition), specificity, F-score, accuracy with a threshold of 0.5 and the area under the curve (AUC) from the receiver operating characteristic (ROC) curve was calculated.

3. Results

The ability of our proposed RRCNN for the classification of DWI data between patients with a low and high DWI-ASPECTS was demonstrated and compared with other deep learning algorithms. Table 2 shows the comparison of results between the proposed RRCNN, the pre-trained VGG16 and Inception V3 model, and the 3DCNN. For the validation data, the accuracy of the proposed RRCNN was 84.4%, which was higher than those of the pre-trained VGG16 (72.8%), pre-trained Inception V3 (72.4%), and 3DCNN (81.7%). Similarly, the AUC of the proposed RRCNN was 0.910, which was higher than those of the pre-trained VGG16 (0.801), pre-trained Inception V3 (0.834), and 3DCNN (0.844). The sensitivity and F1 score also showed a similar trend. The specificity of RRCNN and 3DCNN were similar (89.8% and 89.1%, respectively), which were slightly higher than those of pre-trained VGG16 (86.3%) and pre-trained Inception V3 (84.3%).

Table 2. Comparison of model performances between various deep learning algorithms for classifying patients with high diffusion-weighted magnetic resonance imaging-Alberta Stroke Program Early Computed Tomographic Score (DWI-ASPECTS) (7–10) and low DWI-ASPECTS (1–6).

| Model | Validation Dataset | | | | | Independent Test Dataset | | | | |
|------------------------------------|---------------------------|---------------------------|-------|--------------------------|-------|---------------------------|---------------------------|-------|--------------------------|-------|
| | Sens. ⁴ (%) | Spec. ⁵ (%) | F1 | Acc. ⁶ (%) | AUC | Sens. ⁴ (%) | Spec. ⁵ (%) | F1 | Acc. ⁶ (%) | AUC |
| Pre-trd. ¹ VGG16 | 70.5 | 86.3 | 0.801 | 72.8 | 0.801 | 61.3 | 92.5 | 0.831 | 78.8 | 0.920 |
| Pre-trd. ¹ Inception V3 | 71.8 | 84.3 | 0.795 | 72.4 | 0.834 | 64.5 | 92.5 | 0.840 | 80.0 | 0.921 |
| 3DCNN ² | 78.2 | 89.1 | 0.848 | 81.7 | 0.844 | 77.4 | 90.0 | 0.867 | 84.5 | 0.929 |
| Proposed RRCNN ³ | 82.0 | 89.8 | 0.872 | 84.4 | 0.910 | 83.9 | 90.0 | 0.888 | 87.3 | 0.941 |

¹ Pre-trd., Pre-trained; ² 3DCNN, 3D convolutional neural network; ³ RRCNN, recurrent residual convolutional neural network; ⁴ Sens., Sensitivity; ⁵ Spec., Specificity; ⁶ Acc., Accuracy

Using the independent test dataset, the proposed RRCNN model demonstrated a reasonable performance with a sensitivity, specificity, F1-score, accuracy, and AUC of 83.9%, 90.0%, 0.888, 87.3%, and 0.941, respectively. For all CNN models, the specificity, F1 score, accuracy and AUC of the independent test were either slightly higher than or comparable to those of the validation test. The sensitivities of the RRCNN and 3DCNN were comparable between the validation and independent test datasets, while the sensitivities of pre-trained models in the validation set were higher than those in the independent set. In general, the evaluation metrics showed that the proposed RRCNN had comparable performances with slightly higher levels of sensitivity, specificity, F1 score, accuracy and AUC compared to the other three models. The high level of F1 score (0.888) and the relatively small difference between the sensitivity (83.9%) and the specificity (90.0%) of the RRCNN model indicate that our proposed model did not have a critical bias toward a specific ASPECTS class.

Figure 5a shows the comparison of ROC curves between the four deep learning models. The confusion matrix of the RRCNN model on the independent test data is shown in Figure 5b. Among five cases that were incorrectly predicted as Group 2, three belonged to ASPECTS 6, one ASPECTS 5,

and one ASPECTS 4. Among four cases that were incorrectly predicted as Group 1, two belonged to ASPECTS 7, and the other two to ASPECTS 8.

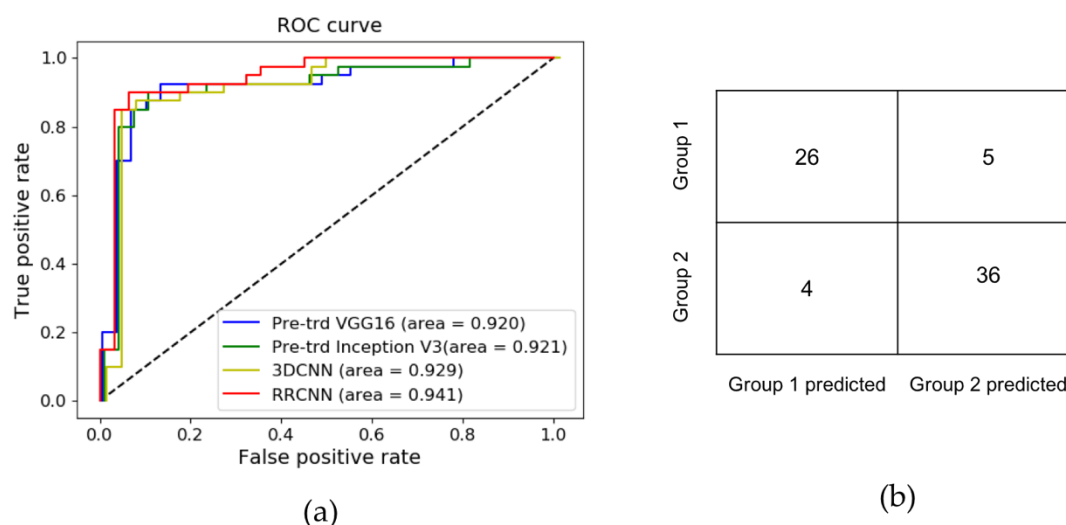


Figure 5. Summary of results using the independent test set. The receiver operating characteristic (ROC) curves are shown for the pre-trained VGG16, pre-trained Inception V3, 3D convolutional neural network (3DCNN), and proposed recurrent residual convolutional neural network (RRCNN) models (a). The proposed RRCNN model demonstrated a slightly higher area under the curve (AUC) (0.941) compared to other models. The confusion matrix demonstrates the classification results using the independent test set (b).

4. Discussion

Estimating the extent of infarction volume plays a pivotal role in the management of patients with acute ischemic stroke and has shown to be an important factor in determining the eligibility of reperfusion therapy and predicting clinical outcomes [4,6,7]. CT-based ASPECTS as well as DWI-ASPECTS are widely used tools for the indirect and rapid assessment of infarction volume [3,5–7]. Several efforts have been made to automate the process of estimating the extent of infarction using computer software. e-ASPECTS is one such method, which was proposed by Hampton-Till et al. and designed for the automated scoring of CT-based ASPECTS [14]. Several studies have demonstrated the feasibility of using e-ASPECTS in assessing CT scans of acute ischemic stroke patients and suggested that the software achieved a performance comparable to stroke physicians or neuroradiologists [22,23]. Other studies have shown that the software can be used to predict the outcome after mechanical thrombectomy [24,25].

Although CT scan is commonly used for the assessment of infarct lesion for stroke patients because of its availability in most emergency clinical environments, the MRI-based assessment of infarction, including DWI-ASPECTS, has shown to provide alternative methods [15,16,26,27]. In a recent study, automated computer-based ASPECT scoring was applied to DWI [28]. The authors utilized a decision tree algorithm to develop an automated method to predict the total ASPECT score. Although the performance of their automated method was slightly worse than human expert scoring, they demonstrated that the machine learning-based method can be used to determine DWI-ASPECTS with good precision. Another automated ASPECTS system combining feature engineering and random forest learning was developed with non-contrast CT scans of 257 patients with acute ischemic stroke using DWI as the ground truth and demonstrated the ability to determine ASPECTS using an automated approach [29]. These studies suggest that machine learning algorithms can be utilized for the automated assessment of ASPECTS.

Deep learning algorithms have been applied to medical imaging [9] and have led to an exciting opportunity for data-driven stroke management and guiding the diagnosis of acute ischemic stroke [30,31]. Recently, several studies have used CNN algorithms for application in acute ischemic lesions

[32–36] and provided effective tools for automatic lesion segmentation or volume calculation. Other studies have focused on developing a deep learning-based approach for detection or identification of large vessel occlusion from CT angiography [37,38]. These studies suggest that deep learning algorithms can be effectively applied to the management of stroke.

Although previous efforts using e-ASPECTS have been shown to provide quantitative scores of ASPECTS, we focused on developing a deep learning algorithm for a fast rendering of a binary classification of DWI-ASPECTS. In ischemic stroke, time is brain, meaning delays to proper treatment lead to worse brain injuries and damage [1,2]. Currently, endovascular treatment is regarded as the standard treatment for selected acute ischemic stroke patients due to large vessel occlusion; however, an appropriate patient selection based on rapid decision-making is important for producing maximal clinical outcome [39]. Previous findings reported marked differences in clinical outcome between two subgroups. The patients with DWI-ASPECTS greater than or equal to 7 had a distinct clinical outcome compared to those with DWI-ASPECTS smaller than 7 after intra-arterial or IV pharmacologic thrombolysis [15,16]. These findings signified the need for determining the different treatment strategies based on the rapid characterization of ischemic lesions between the two groups. The results from our findings suggest that the RRCNN model developed in this study may provide an important ancillary tool for clinicians in a time-sensitive assessment of DWI-ASPECTS from acute ischemic stroke patients. With further improvements and clinical validations, we envision that the algorithm developed in this study may be used for building a triage and fast response system where it can provide a shorter time to mechanical thrombectomy and also a faster time to transfer from a peripheral hospital to a tertiary stroke center, so that the relevant procedure can be performed. This would present a significant clinical benefit because there is a shortage of interventional neuroradiologists [40] and the automated DWI-ASPECTS assessment can help accelerate the procedure of identifying patients who need an interventional neuroradiologist's consultation. Most recently, Viz LVO (Viz.ai, Inc. San Francisco CA, USA), which is software with the similar aim to improve the triage and shorten the time-to-treatment for stroke patients using CT angiography, became the first artificial intelligence-based software to receive the Medicare New Technology Add-on Payment (NTAP) by the Centers for Medicare & Medicaid Services [41,42].

In our previous effort, we provided a similar approach for the rapid assessment of DWI-ASPECTS using a 3DCNN model. Although the previous results demonstrated an accuracy of 81% and AUC of 0.872 for the binary classification DWI-ASPECTS, the 3DCNN possessed more than 100 million parameters, requiring a substantial amount of computation for training. The current study presents several advancements over the previous one. A larger number of data were included in the current study (390 DWI data in the current study vs. 308 in the previous study). In this study, we developed our proposed model based on a recurrent neural network (RNN), which has exhibited promising results in recognizing patterns in sequences of data [43]. In this setting, the multi-slice DWI datasets were regarded as a sequence of images, rather than 3D images. The performance of the proposed RRCNN model was compared to that of the 3DCNN model that has been extensively used for processing 3D datasets [21]. Although the results from the two models were comparable, with our RRCNN model showing a slightly higher level of accuracy (87.3%) and AUC (0.941) compared to the 3DCNN model (accuracy, 84.5%; AUC, 0.929), the computation cost of the proposed RRCNN model was notably reduced compared to that of the 3DCNN, requiring approximately 4 to 5 h of training time, while the 3DCNN required more than 10 to 12 h of training time. These results suggest that RNN-based models may provide an alternative way to analyze multi-slice medical imaging data.

VGG16 [17] and Inception V3 [19] are two of the popular deep neural networks that have been shown to be very efficient in image classification. Although these networks, which were pre-trained with large-scale data from ImageNet [20], have been widely applied to medical image analysis in combination with transfer learning and fine tuning technique [44–48], the effectiveness of the transfer learning method using these deep networks is debatable [49] because medical images, such as MRI and CT, are very different from the images in ImageNet, which are mostly natural images, and the high-level features that are learned during training of medical images can be very different. In addition, the application of transfer learning using these deep networks may need careful

consideration depending on the training condition, such as the number of training datasets, because these complex neural networks generally require a large amount of training data to be effectively trained. Nevertheless, pre-training the well-known networks and applying the transfer learning technique are still very popular methods for medical image analysis. We compared the performance of the proposed RRCNN model that was trained from scratch to those of pre-trained VGG16 and Inception V3. Although this comparison may provide limited information given the data and the specific classification task we applied, the results may be used as a reference regarding the choice of CNN model and training strategy for future studies. The results from our study suggest that the proposed RRCNN model, which has a smaller number of layers and parameters compared to VGG16, Inception V3 and 3DCNN, proved to be as effective, if not more effective, for a rapid assessment of ASPECTS with a limited number of training datasets (Table 2). A recent study reported a similar observation that the pre-training method using images from ImageNet sped up convergence early in training, but did not necessarily provide improved regularization, nor increase test accuracy [49,50].

Although our study demonstrated the potential of the deep learning model to be used for the rapid, automatic assessment of early acute ischemic changes in DWI, it still presents several limitations. First, our study aimed at differentiating low ASPECTS (DWI-ASPECTS 1–6) from high ASPECTS (DWI-ASPECTS 7–10) groups, thereby presenting a global estimation of DWI-ASPECTS rather than a classification of individual DWI-ASPECTS regions. The multi-class classification of individual ASPECTS or a region-based approach may provide an added prognostic value and will be considered in future research. Second, our study included a total of 319 patient DWI datasets acquired from an MR scanner located in our emergency department. Although our results demonstrated the feasibility of using deep learning models for the automatic classification of DWI-ASPECTS, the collection of a larger number of datasets should be considered to improve the performance of the model. In addition, the inclusion of data acquired from MR scanners of different vendors needs to be considered in order to improve the general applicability of our results. A concerted effort to collect data from multiple institutions and validate the developed model with different external datasets is on-going, which is expected to contribute not only to further validate our model, but also to accumulate and share MRI data from stroke patients that can be used in future research. The additional datasets that will be available from the coordinated multi-center collaboration may allow us to develop deep learning models that are able to perform a multi-class classification of DWI-ASPECTS or a region-based analysis of ASPECTS.

5. Conclusions

We developed a deep learning algorithm based on a recurrent residual convolutional neural network for the classification of DWI-ASPECTS. Our model demonstrated an accuracy of 87.3% and AUC of 0.941 for automatic classification between the low and high DWI-ASPECTS. The results suggest that the deep learning algorithm developed in this study can serve as an ancillary tool that assists in the rapid decision making for patients with acute ischemic stroke.

Author Contributions: Conceptualization, L.-N.D., B.H.B., H.-J.Y., I.P.; methodology, L.-N.D., B.H.B.; validation, L.-N.D., B.H.B.; formal analysis, L.-N.D., B.H.B.; investigation, S.K.K., H.-J.Y.; resources, H.-J.Y., I.P., W.Y.; data curation, B.H.B., W.Y.; writing—original draft preparation, L.-N.D., B.H.B.; writing—review and editing, S.K.K., H.-J.Y., I.P., W.Y.; supervision, H.-J.Y., I.P.; funding acquisition, I.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by grants of the Ministry of Education (NRF-2019R1I1A3A01059201), the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (HR20C0021) and Chonnam National University (2020-0362).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Emberson, J.; Lees, K.R.; Lyden, P.; Blackwell, L.; Albers, G.; Bluhmki, E.; Brott, T.; Cohen, G.; Davis, S.; Donnan, G.; et al. Effect of treatment delay, age, and stroke severity on the effects of intravenous

- thrombolysis with alteplase for acute ischaemic stroke: A meta-analysis of individual patient data from randomised trials. *Lancet (Lond. Engl.)* **2014**, *384*, 1929–1935, doi:10.1016/s0140-6736(14)60584-5.
2. Saver, J.L. Time is brain—quantified. *Stroke* **2006**, *37*, 263–266, doi:10.1161/01.STR.0000196957.55928.ab.
 3. Barber, P.A.; Demchuk, A.M.; Zhang, J.; Buchan, A.M.; Group, A.S. Validity and reliability of a quantitative computed tomography score in predicting outcome of hyperacute stroke before thrombolytic therapy. ASPECTS Study Group. Alberta Stroke Programme Early CT Score. *Lancet (Lond. Engl.)* **2000**, *355*, 1670–1674, doi:10.1016/s0140-6736(00)02237-6.
 4. Powers, W.J.; Rabinstein, A.A.; Ackerson, T.; Adeoye, O.M.; Bambakidis, N.C.; Becker, K.; Biller, J.; Brown, M.; Demaerschalk, B.M.; Hoh, B.; et al. 2018 Guidelines for the Early Management of Patients With Acute Ischemic Stroke: A Guideline for Healthcare Professionals From the American Heart Association/American Stroke Association. *Stroke* **2018**, *49*, e46–e110, doi:10.1161/str.0000000000000158.
 5. Grotta, J.C.; Chiu, D.; Lu, M.; Patel, S.; Levine, S.R.; Tilley, B.C.; Brott, T.G.; Haley, E.C., Jr.; Lyden, P.D.; Kothari, R.; et al. Agreement and variability in the interpretation of early CT changes in stroke patients qualifying for intravenous rtPA therapy. *Stroke* **1999**, *30*, 1528–1533, doi:10.1161/01.str.30.8.1528.
 6. Barber, P.A.; Hill, M.D.; Eliasziw, M.; Demchuk, A.M.; Pexman, J.H.; Hudon, M.E.; Tomanek, A.; Frayne, R.; Buchan, A.M. Imaging of the brain in acute ischaemic stroke: Comparison of computed tomography and magnetic resonance diffusion-weighted imaging. *J. Neurol. Neurosurg. Psychiatry* **2005**, *76*, 1528–1533, doi:10.1136/jnnp.2004.059261.
 7. McTaggart, R.A.; Jovin, T.G.; Lansberg, M.G.; Mlynash, M.; Jayaraman, M.V.; Choudhri, O.A.; Inoue, M.; Marks, M.P.; Albers, G.W. Alberta stroke program early computed tomographic scoring performance in a series of patients undergoing computed tomography and MRI: Reader agreement, modality agreement, and outcome prediction. *Stroke* **2015**, *46*, 407–412, doi:10.1161/strokeaha.114.006564.
 8. Lee, D.; Lee, J.; Ko, J.; Yoon, J.; Ryu, K.; Nam, Y. Deep Learning in MR Image Processing. *Investig. Magn. Reson. Imaging* **2019**, *23*, 81–99, doi:10.13104/imri.2019.23.2.81.
 9. Soffer, S.; Ben-Cohen, A.; Shimon, O.; Amitai, M.M.; Greenspan, H.; Klang, E. Convolutional Neural Networks for Radiologic Images: A Radiologist's Guide. *Radiology* **2019**, *290*, 590–606, doi:10.1148/radiol.2018180547.
 10. Yamashita, R.; Nishio, M.; Do, R.K.G.; Togashi, K. Convolutional neural networks: An overview and application in radiology. *Insights Imaging* **2018**, *9*, 611–629, doi:10.1007/s13244-018-0639-9.
 11. Kasasbeh, A.S.; Christensen, S.; Parsons, M.W.; Campbell, B.; Albers, G.W.; Lansberg, M.G. Artificial Neural Network Computer Tomography Perfusion Prediction of Ischemic Core. *Stroke* **2019**, *50*, 1578–1581, doi:10.1161/strokeaha.118.022649.
 12. Winzeck, S.; Mocking, S.J.T.; Bezerra, R.; Bouts, M.; McIntosh, E.C.; Diwan, I.; Garg, P.; Chutinet, A.; Kimberly, W.T.; Copen, W.A.; et al. Ensemble of Convolutional Neural Networks Improves Automated Segmentation of Acute Ischemic Lesions Using Multiparametric Diffusion-Weighted MRI. *AJNR Am. J. Neuroradiol.* **2019**, *40*, 938–945, doi:10.3174/ajnr.A6077.
 13. Do, L.-N.; Park, I.-W.; Yang, H.-J.; Baek, B.-H.; Nam, Y.; Yoon, W. Automatic Assessment of DWI-ASPECTS for Assessment of Acute Ischemic Stroke using 3D Convolutional Neural Network. In Proceedings of the The 6th International Conference on Big Data Applications and Services, Zhengzhou, China, 19–22 August 2018.
 14. Hampton-Till, J.; Harrison, M.; Kühn, A.L.; Anderson, O.; Sinha, D.; Tysoe, S.; Greveson, E.; Papadakis, M.; Grunwald, I.Q. Automated quantification of stroke damage on brain computed tomography scans: E-ASPECTS. *Eur. Med. J.* **2015**, *3*, 69–74.
 15. Aoki, J.; Kimura, K.; Shibasaki, K.; Sakamoto, Y. DWI-ASPECTS as a predictor of dramatic recovery after intravenous recombinant tissue plasminogen activator administration in patients with middle cerebral artery occlusion. *Stroke* **2013**, *44*, 534–537, doi:10.1161/strokeaha.112.675470.
 16. Nezu, T.; Koga, M.; Kimura, K.; Shiokawa, Y.; Nakagawara, J.; Furui, E.; Yamagami, H.; Okada, Y.; Hasegawa, Y.; Kario, K.; et al. Pretreatment ASPECTS on DWI predicts 3-month outcome following rt-PA: SAMURAI rt-PA Registry. *Neurology* **2010**, *75*, 555–561, doi:10.1212/WNL.0b013e3181eccf78.
 17. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* **2014**, arXiv:1409.1556. Available online: <https://arxiv.org/abs/1409.1556> (accessed on 29 September 2020).
 18. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, June 26–July 1 2016.

19. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015.
20. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; et al. ImageNet Large Scale Visual Recognition Challenge. *Int. J. Comput. Vis.* **2015**, *115*, 211–252, doi:10.1007/s11263-015-0816-y.
21. Tran, D.; Bourdev, L.; Fergus, R.; Torresani, L.; Paluri, M. Learning spatiotemporal features with 3D convolutional networks. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 11–18 December 2015.
22. Herweh, C.; Ringleb, P.A.; Rauch, G.; Gerry, S.; Behrens, L.; Mohlenbruch, M.; Gottorf, R.; Richter, D.; Schieber, S.; Nagel, S. Performance of e-ASPECTS software in comparison to that of stroke physicians on assessing CT scans of acute ischemic stroke patients. *Int. J. Stroke Off. J. Int. Stroke Soc.* **2016**, *11*, 438–445, doi:10.1177/1747493016632244.
23. Nagel, S.; Sinha, D.; Day, D.; Reith, W.; Chapot, R.; Papanagiotou, P.; Warburton, E.A.; Guyler, P.; Tysoe, S.; Fassbender, K.; et al. e-ASPECTS software is non-inferior to neuroradiologists in applying the ASPECT score to computed tomography scans of acute ischemic stroke patients. *Int. J. Stroke Off. J. Int. Stroke Soc.* **2017**, *12*, 615–622, doi:10.1177/1747493016681020.
24. Olive-Gadea, M.; Martins, N.; Boned, S.; Carvajal, J.; Moreno, M.J.; Muchada, M.; Molina, C.A.; Tomasello, A.; Ribo, M.; Rubiera, M. Baseline ASPECTS and e-ASPECTS Correlation with Infarct Volume and Functional Outcome in Patients Undergoing Mechanical Thrombectomy. *J. Neuroimaging Off. J. Am. Soc. Neuroimaging* **2019**, *29*, 198–202, doi:10.1111/jon.12564.
25. Pfaff, J.; Herweh, C.; Schieber, S.; Schonenberger, S.; Bosel, J.; Ringleb, P.A.; Mohlenbruch, M.; Bendszus, M.; Nagel, S. e-ASPECTS Correlates with and Is Predictive of Outcome after Mechanical Thrombectomy. *AJNR Am. J. Neuroradiol.* **2017**, *38*, 1594–1599, doi:10.3174/ajnr.A5236.
26. Choi, N.Y.; Park, S.; Lee, C.M.; Ryu, C.-W.; Jahng, G.-H. The Role of Double Inversion Recovery Imaging in Acute Ischemic Stroke. *Investig. Magn. Reson. Imaging* **2019**, *23*, 210–219, doi:10.13104/imri.2019.23.3.210.
27. Park, J.E.; Kim, H.S.; Jung, S.C.; Keupp, J.; Jeong, H.-K.; Kim, S.J. Depiction of Acute Stroke Using 3-Tesla Clinical Amide Proton Transfer Imaging: Saturation Time Optimization Using an in vivo Rat Stroke Model, and a Preliminary Study in Human. *Investig. Magn. Reson. Imaging* **2017**, *21*, 65–70, doi:10.13104/imri.2017.21.2.65.
28. Kellner, E.; Reiser, M.; Kiselev, V.G.; Maurer, C.J.; Urbach, H.; Egger, K. Comparison of automated and visual DWI ASPECTS in acute ischemic stroke. *J. Neuroradiol.* **2019**, *46*, 288–293, doi:10.1016/j.neurad.2019.02.006.
29. Kuang, H.; Najm, M.; Chakraborty, D.; Maraj, N.; Sohn, S.I.; Goyal, M.; Hill, M.D.; Demchuk, A.M.; Menon, B.K.; Qiu, W. Automated ASPECTS on Noncontrast CT Scans in Patients with Acute Ischemic Stroke Using Machine Learning. *AJNR Am. J. Neuroradiol.* **2019**, *40*, 33–38, doi:10.3174/ajnr.A5889.
30. Chilamkurthy, S.; Ghosh, R.; Tanamala, S.; Biviji, M.; Campeau, N.G.; Venugopal, V.K.; Mahajan, V.; Rao, P.; Warier, P. Deep learning algorithms for detection of critical findings in head CT scans: A retrospective study. *Lancet (Lond. Engl.)* **2018**, *392*, 2388–2396, doi:10.1016/s0140-6736(18)31645-3.
31. Feng, R.; Badgeley, M.; Mocco, J.; Oermann, E.K. Deep learning guided stroke management: A review of clinical applications. *J. Neurointerv. Surg.* **2018**, *10*, 358–362, doi:10.1136/neurintsurg-2017-013355.
32. Boldsen, J.K.; Engedal, T.S.; Pedraza, S.; Cho, T.H.; Thomalla, G.; Nighoghossian, N.; Baron, J.C.; Fiehler, J.; Østergaard, L.; Mouridsen, K. Better Diffusion Segmentation in Acute Ischemic Stroke Through Automatic Tree Learning Anomaly Segmentation. *Front. Neuroinform.* **2018**, *12*, 21, doi:10.3389/fninf.2018.00021.
33. Kim, Y.C.; Lee, J.E.; Yu, I.; Song, H.N.; Baek, I.Y.; Seong, J.K.; Jeong, H.G.; Kim, B.J.; Nam, H.S.; Chung, J.W.; et al. Evaluation of Diffusion Lesion Volume Measurements in Acute Ischemic Stroke Using Encoder-Decoder Convolutional Network. *Stroke* **2019**, *50*, 1444–1451, doi:10.1161/strokeaha.118.024261.
34. Sheth, S.A.; Lopez-Rivera, V.; Barman, A.; Grotta, J.C.; Yoo, A.J.; Lee, S.; Inam, M.E.; Savitz, S.I.; Giancardo, L. Machine Learning-Enabled Automated Determination of Acute Ischemic Core From Computed Tomography Angiography. *Stroke* **2019**, *50*, 3093–3100, doi:10.1161/strokeaha.119.026189.
35. Woo, I.; Lee, A.; Jung, S.C.; Lee, H.; Kim, N.; Cho, S.J.; Kim, D.; Lee, J.; Sunwoo, L.; Kang, D.W. Fully Automatic Segmentation of Acute Ischemic Lesions on Diffusion-Weighted Imaging Using Convolutional Neural Networks: Comparison with Conventional Algorithms. *Korean J. Radiol.* **2019**, *20*, 1275–1284, doi:10.3348/kjr.2018.0615.

36. Zhang, R.; Zhao, L.; Lou, W.; Abrigo, J.M.; Mok, V.C.T.; Chu, W.C.W.; Wang, D.; Shi, L. Automatic Segmentation of Acute Ischemic Stroke From DWI Using 3-D Fully Convolutional DenseNets. *IEEE Trans. Med. Imaging* **2018**, *37*, 2149–2160, doi:10.1109/tmi.2018.2821244.
37. Meijs, M.; Meijer, F.J.A.; Prokop, M.; Ginneken, B.V.; Manniesing, R. Image-level detection of arterial occlusions in 4D-CTA of acute stroke patients using deep learning. *Med. Image Anal.* **2020**, *66*, 101810, doi:10.1016/j.media.2020.101810.
38. Öman, O.; Mäkelä, T.; Salli, E.; Savolainen, S.; Kangasniemi, M. 3D convolutional neural networks applied to CT angiography in the detection of acute ischemic stroke. *Eur. Radiol. Exp.* **2019**, *3*, 8, doi:10.1186/s41747-019-0085-6.
39. Wahlgren, N.; Moreira, T.; Michel, P.; Steiner, T.; Jansen, O.; Cognard, C.; Mattle, H.P.; van Zwam, W.; Holmin, S.; Tatlisumak, T.; et al. Mechanical thrombectomy in acute ischemic stroke: Consensus statement by ESO-Karolinska Stroke Update 2014/2015, supported by ESO, ESMINT, ESNR and EAN. *Int. J. Stroke Off. J. Int. Stroke Soc.* **2016**, *11*, 134–147, doi:10.1177/1747493015609778.
40. Davis, S.M.; Campbell, B.C.V.; Donnan, G.A. Endovascular Thrombectomy and Stroke Physicians: Equity, Access, and Standards. *Stroke* **2017**, *48*, 2042–2044, doi:10.1161/strokeaha.117.018208.
41. Viz.ai Granted Medicare New Technology Add-on Payment. Available online: www.prnewswire.com/news-releases/vizai-granted-medicare-new-technology-add-on-payment-301123603.html (accessed on 29 September 2020).
42. Hassan, A.E.; Ringheanu, V.M.; Rabah, R.R.; Preston, L.; Tekle, W.G.; Qureshi, A.I. Early experience utilizing artificial intelligence shows significant reduction in transfer times and length of stay in a hub and spoke model. *Interv. Neuroradiol. J. Perither. Neuroradiol. Surg. Proced. Relat. Neurosci.* **2020**, doi:10.1177/1591019920953055.
43. Sutskever, I.; Vinyals, O.; Le, Q.V. Sequence to Sequence Learning with Neural Networks. In Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014.
44. Chang, J.; Yu, J.; Han, T.; Chang, H.; Park, E. A method for classifying medical images using transfer learning: A pilot study on histopathology of breast cancer. In Proceedings of the 2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom), Dalian, China, 12–15 October 2017.
45. Esteva, A.; Kuprel, B.; Novoa, R.A.; Ko, J.; Swetter, S.M.; Blau, H.M.; Thrun, S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **2017**, *542*, 115–118, doi:10.1038/nature21056.
46. Guan, Q.; Wang, Y.; Ping, B.; Li, D.; Du, J.; Qin, Y.; Lu, H.; Wan, X.; Xiang, J. Deep convolutional neural network VGG-16 model for differential diagnosing of papillary thyroid carcinomas in cytological images: A pilot study. *J. Cancer* **2019**, *10*, 4876–4882, doi:10.7150/jca.28769.
47. Han, S.S.; Park, I.; Lim, W.; Kim, M.S.; Park, G.H.; Chae, J.B.; Huh, C.H.; Chang, S.E.; Na, J.I. Augment Intelligence Dermatology: Deep Neural Networks Empower Medical Professionals in Diagnosing Skin Cancer and Predicting Treatment Options for 134 Skin Disorders. *J. Investig. Dermatol.* **2020**, doi:10.1016/j.jid.2020.01.019.
48. Lakhani, P.; Sundaram, B. Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks. *Radiology* **2017**, *284*, 574–582, doi:10.1148/radiol.2017162326.
49. He, K.; Girshick, R.; Dollar, P. Rethinking ImageNet Pre-Training. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Korea, 27 October–2 November 2019.
50. Shen, Z.; Liu, Z.; Li, J.; Jiang, Y.; Chen, Y.; Xue, X. Object Detection from Scratch with Deep Supervision. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*, 398–412, doi:10.1109/TPAMI.2019.2922181.

