Machine Learning for CT-based Clinical Triage in Hemorrhagic and Ischemic Stroke

D. Kuppens

Thesis submitted for the degree of Master of Science in Technical Medicine

Faculty of Science and Technology
University of Twente
27 September 2018

Student:

D. Kuppens (s1221108)

Research department:

Advanced X-ray Imaging Sciences Center
Department of Radiology
Massachusetts General Hospital

Graduation committee:

Prof. Dr. C.H. Slump - *Committee chair and technological supervisor*
Dr. R. Gupta, M.D. - *Medical supervisor*
Dr. C.O. Tan - *Technological supervisor*
Drs. P.A. Van Katwijk - *Process supervisor*
Dr. B. Sirmacek - *External committee member*

Dedicated to my mother and father, for their lasting love and support.

Summary

Stroke is a 'brain attack'. Stroke occurs when blood flow to an area is disrupted, either by bleeding or by blockage of an artery. To guide management of stroke patients, neuroimaging is essential. The aim of this thesis is to explore to which extent the clinical management of both hemorrhagic and ischemic stroke patients can become more informed by using machine learning techniques to both extract relevant features from more primary data and characterize the state of the patient's stroke.

In intracerebral hemorrhage (ICH), the most occurring subtype of hemorrhagic stroke, computed tomography (CT) accurately visualizes the hematoma. Decisive in treatment of intracerebral hemorrhage is whether there are indicators of further expansion or worse outcome, justifying more invasive treatment. No satisfying marker has yet been found, able to predict hematoma expansion with high sensitivity and specificity. By using dual-energy CT (DECT) angiography, one can quantify the amount of contrast medium in the hematoma and study the texture of the hematoma separately. In this study, we assess the utility of quantitative features of contrast medium distribution and texture analysis of the hematoma in predicting hematoma expansion. Therefore, 15 DECT-based features were designed based on contrast medium quantification in the hematoma and hematoma texture analysis. The best performing combination of features was modeled using a support vector machine with radial kernel to classify hematomas as either expanding or stable. On the train set of 45 hematomas, our model outperformed the conventional spot sign in our cohort and in literature on sensitivity and specificity. On the test set of 69 hematomas, our model outperformed the conventional spot sign in our cohort and in literature on specificity, but not on sensitivity. Although the drop in sensitivity of our model between the train and test set indicates that more data and a more standardized image acquisition is required to make a model with a higher generalization ability, we believe that quantitative features can aid in the prediction of hematoma expansion.

In acute ischemic stroke (AIS), treatment is based on the size of the ischemic lesion and on clinical presentation. CT is not able to quantify the volume of the ischemic lesion in the acute setting. However, the quality of the collateral circulation and the quality of tissue perfusion, related to the volume of the ischemic lesion, can be estimated using CT angiography and delayed-phase CT angiography. In this study, we explore the use of deep learning to predict lesion shape and size using multiple CT acquisitions. We have designed a deep convolutional neural network (CNN), transforming a CT input into a probability distribution of the pixels belonging to an ischemic lesion. The network was trained on a train set of 53 patients, minimizing the Dice loss between the probability distribution, and a segmented mask of the ischemic lesion on diffusion weighted imaging (DWI). On the train set, the network reached a Dice loss <0.4. Unfortunately, on the validation set the Dice loss remained close to 1. This indicates that our network overfitted on the training data, and therefore was not able to generalize to new, unseen data.

We demonstrate that machine learning can be useful for DECT-based clinical decision making in ICH triage. In AIS triage, our deep CNN is not yet able to predict ischemic lesion shape and size based on CT acquisitions.

Contents

Chapter 1

Introduction

1.1 Stroke

Each year, almost 800.000 people suffer from stroke, leading to about 1 of every 20 deaths in the United States [1]. Stroke is caused by a disruption of blood flow to the brain, depriving brain tissue of oxygen and other essential nutrients. Stroke can be caused by either a bleeding artery, called hemorrhagic stroke, or a blocked artery, called ischemic stroke. Hemorrhagic stroke can be further categorized as either intracerebral hemorrhage (ICH), bleeding within the brain tissue or ventricles, or subarachnoid hemorrhage, bleeding in subarachnoid space (SAH). 87% of all strokes are ischemic strokes, while only 10% are caused by ICH and 3% are caused by SAH [1, 2].

1.1.1 Intracerebral hemorrhage

Pathophysiology

ICH is an acute and spontaneous bleeding into the brain parenchyma, resulting from spontaneous rupture of small arteries [3]. The 30-day mortality after initial ICH is reported to be as high as 50%, with approximately half of all deaths occurring in the first 48 hours after treatment [4]. The most common risk factors for primary ICH are hypertension and cerebral amyloid angiopathy, both accelerating vascular deterioration [5, 6]. In deep hematomas, spontaneous vessel rupture is attributed to degenerative changes in the vessel wall [7]. In ICH related to cerebral amyloid angiopathy, the cause of vessel rupture is a combination of amyloid deposition and breakdown of the vessel wall [8]. ICH can be a multiphasic event, with persistent bleeding and resulting expansion of the hematoma up to 6 hours in noncoagulopathic ICH and up to 24 hours in coagulopathic ICH [9]. Hematoma expansion is a common and serious complication after ICH, potentially leading to hydrocephalus, severe tissue shifts, increased intracranial pressure and mass effect [10], visualized in fig. 1.2. Hematoma expansion is strongly associated with increased mortality, worse functional outcome and neurologic deterioration [12, 13, 14]. Each millimeter of hematoma expansion is associated with an increase of risk on long-term disability of 7% [14]. The mechanism driving hematoma expansion has not been fully explained, possibly being caused by the inflammatory response after initial bleeding, hemostatic dysregulation, breakdown of the blood-brain barrier or local tissue distortion [15, 16, 17].

Treatment

Due to the risk of rapid deterioration shortly after ICH onset, a correct diagnosis is essential to guide quick and adequate treatment. Relief of secondary symptoms is an essential part of ICH treatment, since hematoma expansion increases risk on rapid neurologic deterioration and

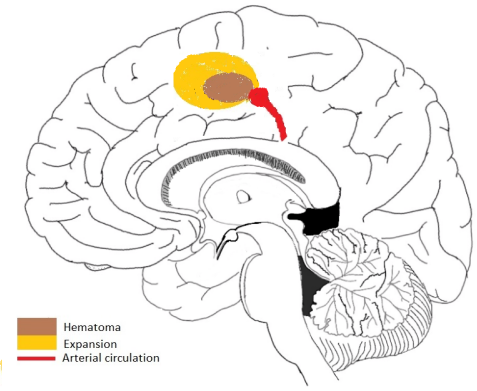

Figure 1.1: Schematic representation of spontaneous intracerebral hemorrhage and possible expansion of the hematoma. Adapted from [11].

mortality. Hemostatic agents such as recombinant factor VII have shown to significantly reduce hematoma volume growth, but fail to reduce mortality or improve functional outcome. The lack of improvement in clinical outcome may be due to a reported increase in serious arterial thromboembolic adverse events after recombinant factor VII administration [18]. If and when surgical intervention should be done in ICH remains an ongoing topic of discussion. While some studies show a tendency for improved immediate outcomes after craniectomy in ICH, a long-term improvement in outcomes has not been demonstrated compared to medical management only [19, 20, 21, 22]. However, neurosurgical intervention is standard of care if the intracerebral hemorrhage is large or if it causes intraventricular shift, hydrocephalus, obtundation or brainstem compression [10]. Fig. 1.6 shows the protocol at our institution for imaging-based decisions on treatment of ICH.

Imaging

Abrupt onset of focal neurological symptoms has to be presumed to be of vascular origin until proven otherwise. However, due to the lack of hemorrhage-specific clinical symptoms, discrimination between acute ischemic stroke or hemorrhage on clinical characteristics alone is difficult [23]. For initial imaging in the evaluation of focal neurological symptoms with abrupt onset,

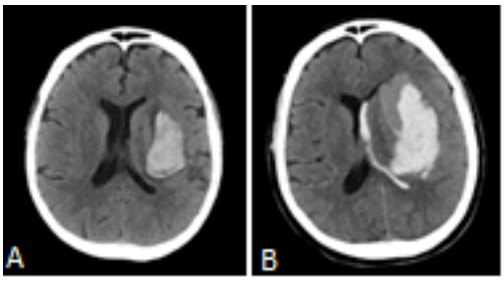

Figure 1.2: CT imaging of ICH. A: Parenchymal intracerebral hemotoma appears hyperdense on non-contrast CT. B: Hematoma expansion on follow-up non-contrast CT.

CT is sensitive in the identification of acute hemorrhage and is considered the 'gold standard'. Although magnetic resonance imaging (MRI) acquisitions such as gradient echo and T2* susceptibility weighted imaging are equally sensitive, the time, cost, availability and patient tolerance issues involved with acute MRI prevent the use of magnetic resonance (MR) in most cases [24, 25]. CT angiography has been used as a means to characterize risk of hematoma expansion in acute presentations. 28% to 38% of patients undergoing head CT within 3 hours of ICH onset are reported to have hematoma expansion [26, 14]. Therefore, timely identification of patients at risk for hematoma expansion is critical in management of ICH. Non-radiographic risk factors for hematoma expansion include early presentation [27, 28, 29] and antecedent use of anti-coagulation medication [30, 31]. Several studies also found that high systolic blood pressure is a risk factor for hematoma expansion [32, 33, 34], while others did not [35]. Although many radiographic markers have been used to qualitatively assess ICH on both non-contrast computed tomography (NCCT) and computed tomography angiography (CTA), they lack adequate sensitivity in general [36, 37]. In NCCT, examples of such markers are hematoma location [38] and markers that describe morphologic appearance of a hematoma, such as shape and density variation, possibly reflecting active hemorrhage, more variable hemorrhagic time course or multifocality of bleeding, which are both considered as independent predictors of hematoma expansion [35, 39]. Specific radiographic markers taking into account shape or density variations in the hematoma are: the island sign [40], satellite sign [41], black hole sign [42], blend sign [43], swirl sign [44], margin irregularity [45] and hematoma hypodensities [46]. In CTA, the presence of foci of hyperintensities in the hematoma - known as the spot sign and assumed to represent active contrast extravasation - is independently associated with an increased risk of hematoma expansion and has a specificity of $80-90\%$ and a sensitivity of $50-60\%$ in predicting hematoma expansion [47, 48, 49]. Although delayed CTA imaging is able to improve spot sign sensitivity, triage based on spot sign presence on delayed CTA did not improve treatment efficacy [50, 51]. So far, none of all these markers are able to show good predictive value of identifying patients who will likely develop hematoma expansion with both high sensitivity and specificity.

1.1.2 Acute ischemic stroke

Pathophysiology

Acute ischemic stroke (AIS) is a focal neurological condition with sudden decrease of blood supply to an area of the brain. A reduction of the cerebral blood flow to below a threshold value leads to a series of functional, biochemical and structural changes and eventually to irreversible neuronal death if not reperfused quickly enough [52]. Affected brain tissue can be divided into ischemic core, penumbra and benign oligemia according to the severity of the ischemia, as presented in fig. 1.3. The ischemic core is an area of irreversibly injured brain tissue at the time of imaging. Surrounding the core lies a penumbra which is functionally impaired but still viable hypoperfused tissue that may be recruited into the infarct core if perfusion is not restored on short term [53, 54, 55, 56]. The penumbra is predominantly maintained by blood supply from the collateral circulation, meaning the rate of progression into infarction without reperfusion is dependent on the quality of this collateral circulation and the duration of the insult [52, 57], as visualized in fig. 1.4. The penumbra can be surrounded by benign oligemia, which is hypoperfused tissue with normal function, recovering spontaneously irrespective of improvement in blood supply [52, 58].

Treatment

The purpose of therapy in AIS is to save the penumbra from culminating into infarction. Intravenous alteplase (IV-tPA) administered within the first 4.5 hours has long been the only reperfusion therapy with proven efficacy in patients with AIS [59]. However, the limited efficacy of

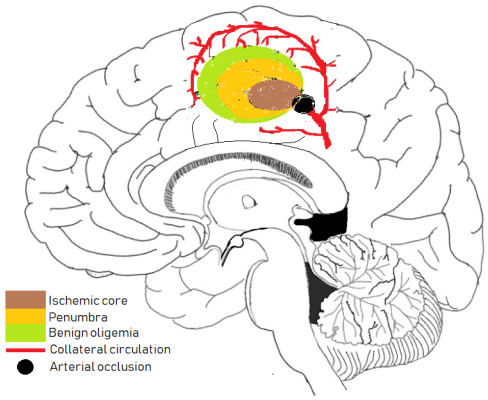

Figure 1.3: Representation of ischemic core (non-viable tissue), penumbra (hypoperfused tissue which will be recruited into ischemic core without timely reperfusion) and benign oligemia (hypoperfused tissue which will recover independent of reperfusion). Collaterals supply the benign oligemia and penumbra with blood in case of occlusion of the irrigating artery. Adapted from [11].

IV-tPA was due to a moderate rate of early reperfusion in patients with an LVO [60]. 60-80% of patients with a proximal vessel occlusion in the anterior circulation treated with IV-tPA die or lack functional independence 90 days after stroke onset [59, 61]. In addition, because of the strict therapeutic time window and a high risk for cerebral and systemic hemorrhage, less than 10% of patients presenting with ischemic stroke met the eligibility criteria for the use of IV-tPA [62, 63]. Because of the multitude of constraints involved in IV-tPA treatment, trials studying the clinical efficacy of intraarterial treatment (IAT) have been set up. Intraarterial treatment can be divided into intraarterial administration of alteplase and endovascular thrombectomy (EVT). Initial trials studying the effects of IAT in combination with IV-tPA compared to IV-tPA alone, such as SYNTHESIS, IMS III & MR RESCUE, failed to demonstrate a beneficial clinical effect. However, these trials were limited by the intraarterial administration of t-PA or endovascular thrombectomy using early-generation thrombectomy devices, a long period between onset of stroke symptoms and start of IAT and a lack of adequate vessel imaging to confirm the patients' eligibility for IAT [61, 64, 65]. Subsequently, the MR CLEAN trial has been able to demonstrate a beneficial clinical effect of IAT in selected patients by employing more strict inclusion criteria, using third-generation mechanical thrombectomy devices and ensuring imaging-confirmed occlusion of the anterior circulation. Results of the MR CLEAN trial shifted focus of triage in AIS to selecting the specific subset of patients who are expected to have most clinical benefit from EVT. In selected patients, based on qualitative CTA-based assessment of the quality of the collateral circulation [60] or the CT perfusion derived size of the ischemic core and penumbra [66, 67], EVT combined with IV-tPA administration has been shown to have an even larger clinical benefit versus IV-tPA administration alone in patients with AIS when performed within 6 hours after stroke onset. Recently, the DAWN trial showed that even patients who undergo treatment 6-24 hours after stroke onset and have a mismatch between clinical deficit and infarct benefit from EVT combined with standard care compared to standard care alone, if selected correctly [68]. The evolution of the demonstrated clinical effect of EVT indicates that EVT will have most clinical benefit in selected patients with an imaging-confirmed proximal occlusion of the intracranial circulation with a large area of salvageable brain tissue and a small ischemic core [59, 67, 60, 66, 69, 68]. Several studies have shown the decisive role of infarct core size in predict-

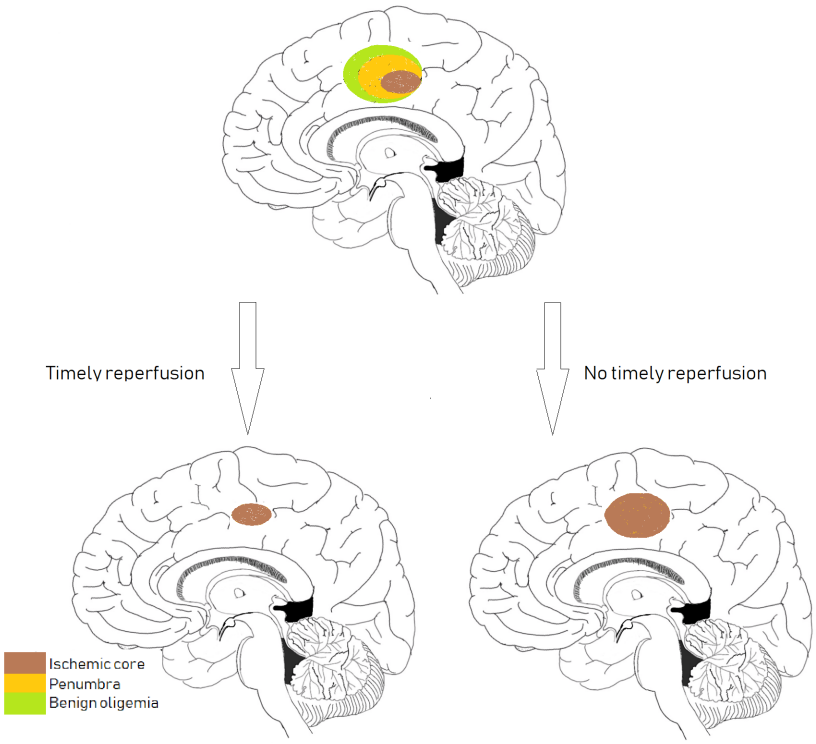

Figure 1.4: Evolution of the infarcted tissue depending on the time before reperfusion is established. If reperfusion is established quickly, the ischemic core will recruit only minimal portions of the penumbra. Without timely reperfusion, the penumbra will be entirely recruited into the infarct core. Adapted from [11].

ing the long-term functional outcome after EVT [70, 71]. Patients presenting with large infarcts have a small chance for a beneficial response to IAT [72, 73], and a higher risk of reperfusion hemorrhage [74, 75]. Thus, patients with a large infarct at baseline imaging will most probably not be eligible for EVT.

Imaging

Patients with a treatable occlusion confirmed on imaging therefore depend on infarct size to determine eligibility for EVT, and imaging should quickly, accurately and reliable define infarct core. The triage of AIS patients is based on three major imaging features: exclusion of intracranial hemorrhage and stroke-like mimics, detection of the site of arterial occlusion and the determination of the extent of the ischemic lesion. For the initial assessment in AIS triage, imaging should exclude the possibility that symptoms are caused by brain hemorrhage or stroke mimics. Both CT and MRI have are considered gold standard for detection of brain hemorrhage [76, 77]. If brain hemorrhage and stroke mimics are excluded, vessel imaging is performed to evaluate the site of arterial occlusion. AIS caused by proximal large vessel occlusion is better accessible for thrombectomy and less likely to recanalize by IV-tPA administration alone and can therefore be considered for EVT [78]. To evaluate the brain parenchyma for the extent of infarction in AIS, diffusion weighted imaging (DWI) and the apparent diffusion coefficient (ADC) maps are the gold standard for identification of infarct core in AIS, being the most reliable and most accepted

technique to measure infarct size [56, 79, 78, 80, 81]. Even though the delineation of infarct core

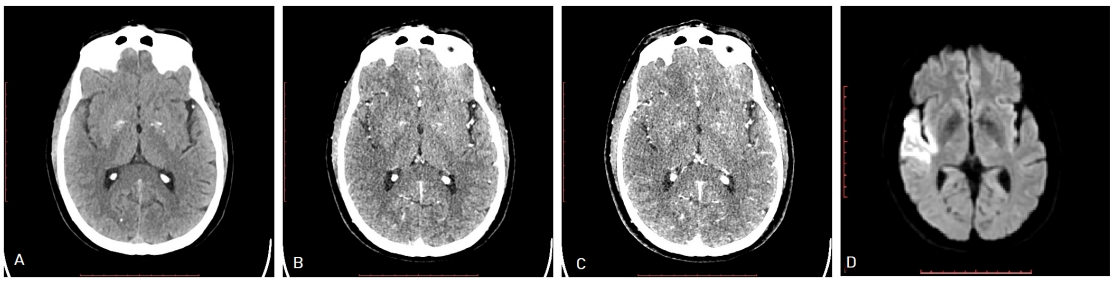

Figure 1.5: Data present after imaging with the Massachusetts General Hospital acute ischemic stroke protocol. A: non-contrast CT, B: dual-energy CT angiography 80 kVp, C: dual-energy CT angiography 140 kVp, D: DWI MR scan. AIS shows increasing loss of grey matter - white matter differentiation over time on non-contrast CT. In case of ongoing occlusion, a discontinuity in hyperdense arterial vessels caused by presence of contrast medium can be seen on CTA. On DWI, ischemic lesions are hyperintense.

is not perfect, DWI is able to visualize AIS with a sensitivity of up to 73-92% within 3 hours after onset of stroke symptoms and up to 95-100% within 6 hours [76, 82, 83]. For detection of acute infarcted areas, CT is significantly less sensitive than DWI with a sensitivity of up to 12% in the first 3 hours after onset of stroke symptoms and up to 57-71% in the first 24 hours. Perfusion imaging has been used in several trials to evaluate cerebral hemodynamics as a measure of identification of ischemic core and penumbra [67, 66, 68]. However, while perfusion imaging assesses the cerebral hemodynamics at a single point in time, there is no single hemodynamic state that uniquely and robustly characterizes the infarct core [56]. CT perfusion has a high burden of radiation exposure and suffered from a lack of clear guidelines for indication, acquisition methods and interpretation leading to a high variability in data [56]. As an alternative to DWI and CT perfusion imaging, AIS patients have been successfully selected for EVT based on multiphase CT angiography to qualitatively assess collateral function, based on filling of the pial arterial circulation of the middle cerebral artery (MCA) [60]. In addition to traditional performance metrics, in the case of AIS other factors also weigh in on the decision to perform imaging for patient selection, such as optimal triaging workflow and availability, reliability and repeatability of the imaging modality [56]. Several trials have shown successful patient selection based on different triage workflows. The common denominator in nearly all workflows is the presence of a non-contrast CT scan for the exclusion of intracranial hemorrhage and stroke like mimics and the acquisition of CT angiography to confirm the site of large vessel occlusion [67, 66, 68, 69, 60]. Upon comparison of AIS patients treated with EVT between different methods of patient selection, all with similar times from stroke onset to EVT puncture, the rates of patients with functional independence after EVT range from 53% to 71% and the ratio of patients treated to patients screened varies from 1:3 to 1:14 [69, 60, 67, 66]. Patient selection for EVT in AIS should optimize the trade-off between clinical efficacy in the treated group and the ratio between treated and screened patients. Fig. 1.6 shows the patient selection protocol at our institution, selecting only those patients with non-large infarct volumes on DWI, and is estimated to treat 1 in every 3 screened patients, while reaching functional independence at 90 days in 53% of the patients in AIS patients treated with EVT [56, 69]. This DWI-based EVT patient selection might be close to the optimal trade-off between clinical efficacy and number of patients treated [69]. However, the cost and workflow related constraints involved in implementing MRI in an acute setting raise the question if the patient selection can be performed using less resources.

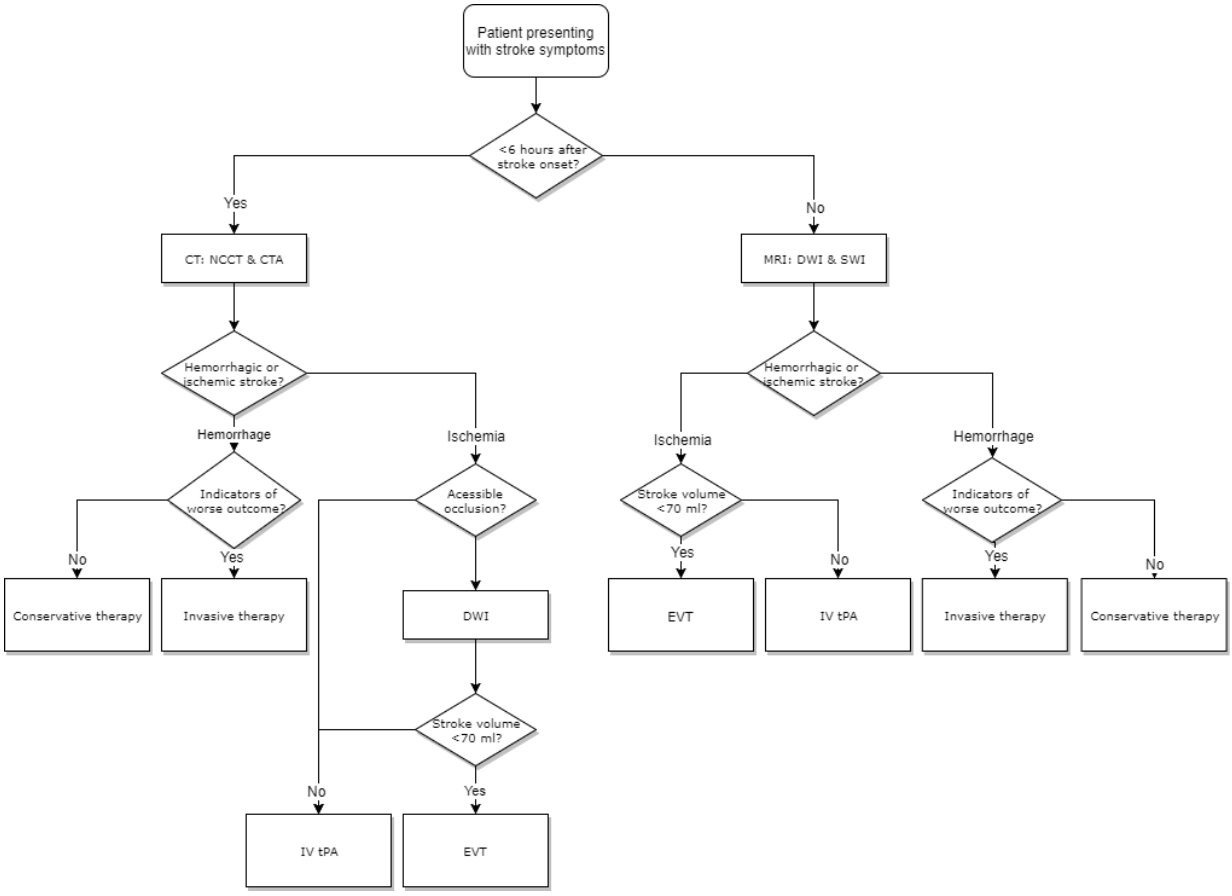

Figure 1.6: Stroke imaging protocol implemented at the Massachusetts General Hospital. If patients present within 6 hours of onset of stroke symptoms, primary neuroimaging is CT-based, otherwise primary neuroimaging is MRI-based. Primary neuroimaging has the objective to assess the etiology of the stroke symptoms: hemorrhage, ischemia, or stroke-like mimics. In case of ischemic stroke, patients are selected for EVT based on the extent of the ischemic lesion. In case of hemorrhagic stroke, patients are selected for invasive therapy based on the presence of spot sign or other indicators of worse outcome. Adapted from the Massachusetts General Hospital Stroke protocol [56].

Collateral function

The quality of the collateral function is an independent predictor of stroke outcome, regardless of treatment, and a small lesion on follow-up imaging. Results from the ESCAPE, EXTEND-IA, SWIFT PRIME and DAWN trial demonstrated that the patients who will benefit most from EVT within 6 hours and within 24 hours after onset of stroke symptoms are patients with a small ischemic core and large amount of salvageable brain tissue [60, 67, 66, 68]. In addition, it has been shown to predict response to IV-tPA [84, 85, 86]. Good quality collateral function is associated with large penumbra size and a small infarct core size at presentation, where worse collateral function is associated with rapid infarct progression [85, 86]. Absence of collaterals at delayed phase CT angiography is a specific predictor for large DWI infarct core at presentation and poor clinical outcome [87]. The quality of the collateral circulation can be measured by imaging of

retrograde filling of arteries distal to the occlusion using delayed-phased CT angiography, digital subtraction angiography (DSA), dynamic CT angiography (dCTA), contrast-enhanced MR angiography (CE-MRA), and dynamic MR angiography (dMRA). Other than invasive modalities (DSA) or modalities constrained by availability (MR), CT angiography is already incorporated in the AIS imaging workflow, widely available, non-invasive and able to qualitatively assess retrograde filling and laterality in parenchymal enhancement [57, 60]. CT techniques combining acquisitions over multiple time-points are able to assess collateral flow despite variations in cerebral hemodynamics [88, 89]. Several studies for evaluating the function of the collateral circulation in AIS patients have been performed based on qualitative observations on 2-dimensional CT angiography maximum intensity projections (CTA MIP) and 3-dimensional CT angiography source images (CTA-SI) based scoring systems of the affected vascular territory relative to non-affected brain territories [90, 86]. Collateral function can also be assessed by the level of hypoattenuated brain tissue in CT angiography, since hypoattenuation in CT angiography does not only reflect the tissue density as in non-contrast CT, but also incorporates the amount of contrast received in the brain tissue [91]. A typical CT angiography measurement is shown in fig. 1.5.

1.2 Machine learning

Medical imaging has become indispensable in the diagnosis and therapy of diseases. Advances in medical imaging increase the amount of information available for making an informed clinical decision. To make use of all available information in medical imaging, technologies to extract the information from raw data and transform relevant information into a clinical decision have to be developed. Machine learning identifies patterns in existing data representations and uses these identified patterns to make predictions on new, unseen data. Supervised machine learning seeks to learn a model, $f(\mathbf{x}) = y$, returning an output state, $\widehat{y}$, according to a conditional distribution function, $P(y|\mathbf{x})$, based on N known data pairs $(\mathbf{x}_i, y_i)$, with $i = 1, \ldots, n$, composed of feature vectors $\mathbf{x}$, with each feature vector consisting of M features with $\mathbf{x}_i \in \mathbb{R}^m$ and output states y, see fig. 1.7. Since the objective is to define a model able to generalize to new data, it is

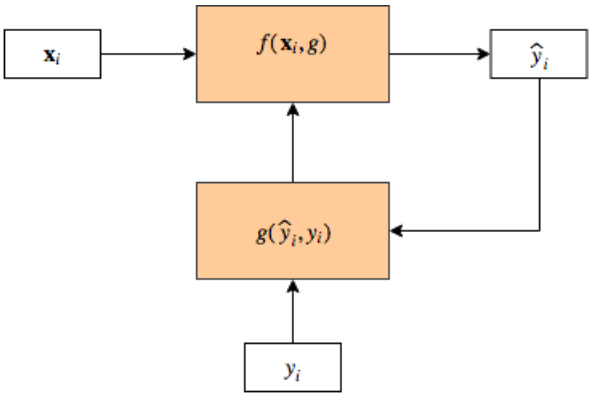

Figure 1.7: A predictive model, f, estimates the mapping of a feature vector, $\mathbf{x}_i$, to an output variable, y_i. The mapping iteratively improves by adapting to minimize the discrepancy between the estimated output, $\widehat{y}_i$, and the supervision, y_i, using a loss function, g. When a new, unseen feature vector is presented, the predictive model predicts the corresponding output variable.

assumed that the data pairs $(\mathbf{x}_i, y_i)$ are drawn independently from a fixed, unknown probability distribution $P(\mathbf{x}, y)$, which is identical for all feature vectors $\mathbf{x}$. Without putting restrictions on

our model, f could do particularly well on training data but need not generalize well to new data. A small training error does not imply a small test error. Therefore, the model should be restricted to the set of functions f suitable for the amount of training data available [92, 93].

1.2.1 Support vector machine

In the case of classification of hematomas in expanding and stable hematomas, the objective is to separate the feature vector $\mathbf{x}_i = x_{i,1}, x_{i,2}, \ldots, x_{i,m}$ in two classes, $y = \pm 1$, using a classification hyperplane T. If the assumption of normality regarding the distribution of $\mathbf{x}$ is violated, traditional methods to determine the optimal decision boundary overemphasize outliers, thereby negatively influencing the classification boundary. A distribution-free method to decrease the influence of outliers is the use of a support vector machine (SVM). SVMs construct a classification hyperplane in feature space based on the position of the data samples close to data samples of the opposite class, the support vectors [94]. By constructing the decision boundary based on these data samples, the margin between the two classes is maximized. In a linearly separable case, the classification hyperplane separating feature vectors in two classes is:

$$\mathbf{w}^T\mathbf{x} + b \begin{cases} < 0 & \text{if } y = -1, \\ \geq 0 & \text{if } y = 1. \end{cases} \tag{1.1}$$

With weight vector $\mathbf{w}$, feature vector $\mathbf{x}$, bias b and class y. If the classes are not linearly separable, the original feature vector, $\mathbf{x}$, is mapped to a feature vector in a higher-dimensional feature space, $\phi(\mathbf{x})$, using a non-linear transformation ϕ, to improve linear separation of classes, see fig. 1.8.

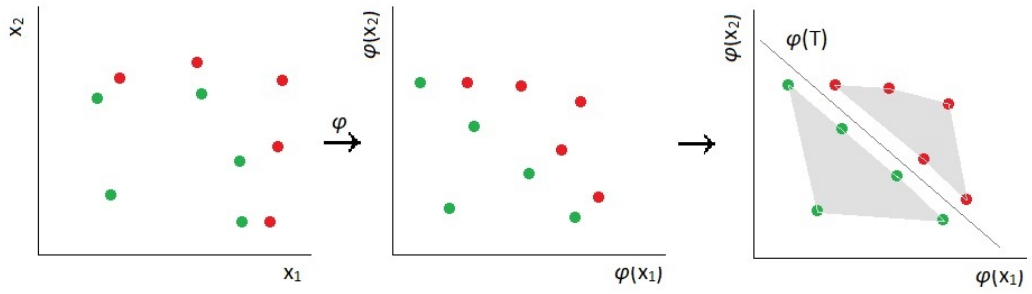

Figure 1.8: Non-linear mapping, ϕ, of the feature vector, $\mathbf{x}_i$, into higher-dimension feature space where the feature vector, $\phi(\mathbf{x}_i)$ becomes easier to separate linearly. A linear classification boundary is drawn perpendicular to the line between the closest points on the two hull convexes surrounding the data samples of each class in feature space.

Kernel function

In the maximization of the classification margin, feature vectors only show up in the form of a mutual dot product. In maximization of the classification margin in feature space, the dot product between two feature vectors, $(\mathbf{x}_i \cdot \mathbf{x}_j)$, is replaced by the dot product of those feature vectors in feature space $(\phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_j))$. To avoid computation of the feature vector in feature space, implementation of a kernel function allows the computation of the equivalent of the classification hyperplane in feature space, without explicitly computing the feature vector in feature space [94].

The kernel function, K, defines dot products in feature space:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_j). \tag{1.2}$$

1.2.2 Deep learning

In clinical practice, detection and diagnosis of diseases using medical imaging is predominantly performed by trained humans. In the case of AIS, a radiologist determines whether a patient is eligible for EVT based on lesion volume on DWI, core volume on a perfusion CT or collateral luminance on a CT angiography. Human analysis of ischemic lesion extent on both DWI and CT perfusion is quantitative and accurate. However, quantification of collateral function is hard

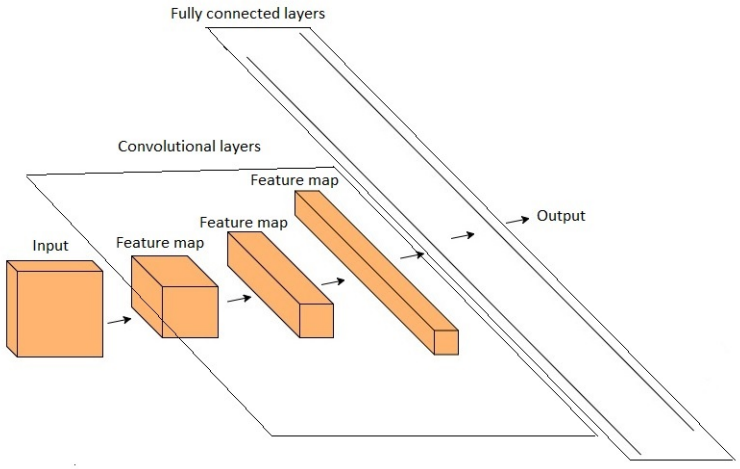

Figure 1.9: Representation of a convolutional neural network. A convolutional neural network transforms the input via feature maps in several convolutional layers, consisting of convolution operations with a kernel composed of learned weights, a non-linear activation function and a subsampling operation and fully connected layers into the output. Each convolutional layer transforms the input in a non-linear way, decreasing image dimensions, increasing feature map quantity and increasing the complexity that can be captured in the feature maps. In a typical CNN, the supervision is a classification.

based on the diffuse presence of collaterals, the small volume of collaterals and the influence of delay time on collateral luminance. In machine learning, patterns in informative features present within sample data are learned to make generalizable data-driven models for prediction and classification of new data samples. Conventional machine learning techniques are limited in their capability to process raw data and require engineering of specifically designed features to transform raw data into a meaningful representation from which patterns can be recognized. In representation learning, the manual feature engineering step is replaced by automated learning of meaningful data representations - features - inherent in observed raw data samples. Deep learning is a set of representation learning methods characterized by multiple, hierarchical layers of data representations of increasing complexity, achieved by performing non-linear operations at each layer. The distinguishing principle behind deep learning is that these abstract features are learned instead of human-designed.

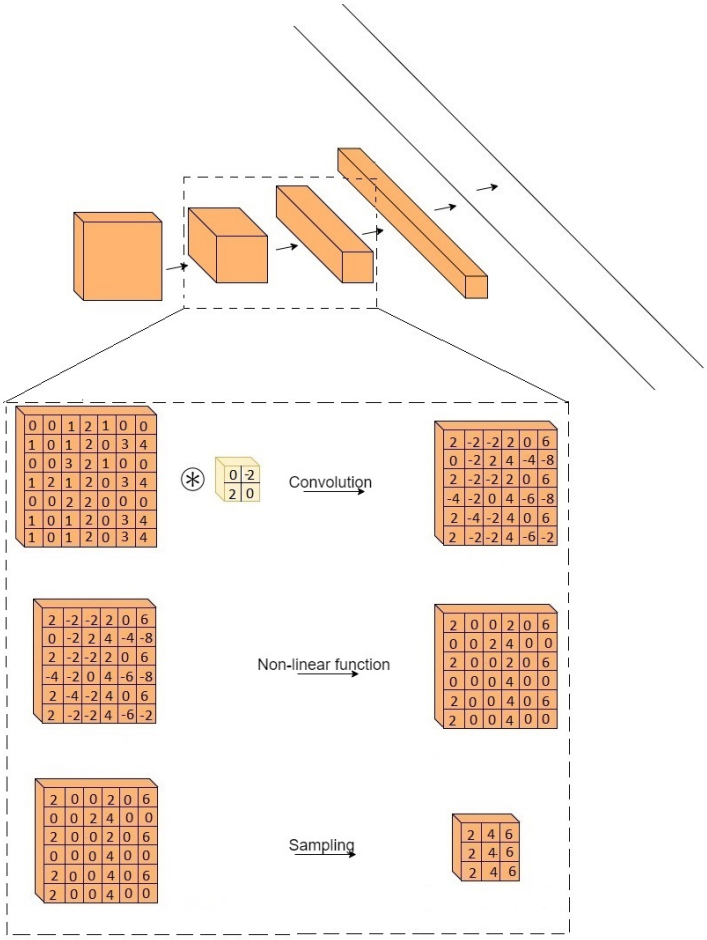

Figure 1.10: Operations in a convolution layer: convolution, non-linear operation and sampling. The top row shows convolution of a layer input with a kernel composed of learned weights using a stride of 1 in both directions and no padding. The middle row shows the effect of a non-linear operation, a rectified non-linear unit, on the result of the convolution operation. The bottom row shows a max-pooling subsampling operation with stride 2 in both directions on the non-linearly activated convolved input.

Convolutional neural networks

In medical image analysis, deep learning models have to make optimal use of the structural information present in image data, must be insensitive to irrelevant variations in the input data, such as variations in translation, rotation, scaling and illumination, and very sensitive to relevant variations in the input data, such as variations in luminance of the collaterals in AIS detection. Convolutional neural networks (CNN) are deep, feed-forward artificial neural networks specifically designed to make optimal use of the spatial configuration of data that come in the form of tensors. A typical CNN is visualized in fig. 1.9. A CNN is structured as a series of hierarchical layers. Deep CNNs allow the detection of complex features by composing higher level, abstract features from low-level features. A CNN utilizes spatial properties in tensors using four concepts: multiple layers, sampling, shared weights and local connections [95]. The principal building block

of a CNN is the convolutional layer, detecting local features in the input by performing convolution operations with a kernel composed of learnable weights on spatially contiguous subsets of the input. A convolutional layer consists of feature maps, in which each feature map is connected to a feature map of the previous layer by a set of learnable weights in the convolution kernel. Every unit in one feature map has the same set of weights, but different feature maps use different sets of weights. The shared weights in each feature map make a convolutional layer spatially invariant, detecting equal features on different positions in the input tensor. On the output of the convolution operation a non-linear operation is performed, a combination of multiple non-linear operations layers throughout the various network layers yields a more complex mapping from network input to output. Sampling the non-linear convolution output removes less prominent features by only taking one unit from a group of values, thereby reducing image dimensions and making the network invariant to small shifts and differences in configuration in each layer. The convolution operation, non-linear activation and sampling are visualized in fig. 1.10. Conventionally, CNNs are used for classification tasks, assigning a label to an input tensor (e.g. an image). However, in characterization of stroke lesions, the desired output has a structured form. Fully convolutional networks adapt the conventional CNN architecture and replace fully connected layers by deconvolution layers, upsampling the sparse representation to the original input resolution while decreasing the amount of feature maps per layer [96].

Training convolutional neural networks

The key advantage of deep learning is that the extraction of features can be learned autonomously using a general-purpose learning procedure. All functional elements composing a convolutional neural network are subject to learning. Supervised learning is the most common method to train a deep CNN, where each input has a corresponding supervision reflecting the desired output. An objective cost function, $g(\widehat{\mathbf{y}}, \mathbf{y})$, compares the distance between the network output, $\widehat{\mathbf{y}}$, and the supervision over a batch of training data, $\mathbf{y}$. During training, weights are adjusted iteratively to minimize g. The estimation of the impact of small variations in weights is measured by the gradient of the loss function with respect to the weights. The set of weights, $\mathbf{W}$, is iteratively adjusted using a procedure called stochastic gradient descent (SGD) :

$$\mathbf{W} = \mathbf{W} - \eta \nabla g(\mathbf{W}). \tag{1.3}$$

Where $g(\mathbf{W})$ is the loss over a mini-batch as a function of the set of weights as estimation of the $g(\mathbf{W})$ over the entire train set and η is the learning rate [97]. Starting at the output of the cost function, the gradient vector backpropagates through the layers of the network. At each layer, the gradient for each weight kernel can be computed using the chain rule. The weight in the convolution kernel is corrected in the opposite direction of the gradient. The weight optimization continues until the average of the cost function over a batch of training data stops decreasing, meaning a local minimum for the cost function in multi-dimensional weight space has been reached. After training, the kernel weights are fixed and the model performance is tested on an independent test set. The performance on the test set compared to the performance on the train set gives an indication of the model's ability to generalize.

Chapter 2

Research objective

Stroke is one of the most devastating medical conditions. Both hemorrhagic stroke and ischemic stroke are caused by disruption of blood flow to the brain, either due to bleeding or blockage of an artery and require a quick judgment on the type of treatment required. Too aggressive management of both hemorrhagic and ischemic stroke patients induces risk for the development of serious adverse events. Endovascular thrombectomy in AIS patients with a stroke too large creates a high risk for hemorrhagic conversion, restraining therapy from being beneficial. Conversely, too aggressive medical therapy in ICH patients leads to a high risk for arterial thromboembolic events. Since aggressive therapy on the entire population of ICH or AIS patients has not been proven to be beneficial, stratification of stroke patients for risk of serious adverse events in aggressive therapy may improve the long-term outcomes of therapy. In AIS triage, the risk of serious adverse events during or after EVT is too high when the ischemic lesion is too large, therefore the objective is to detect patients with an ischemic lesion on CT, and to be able to differentiate between large ischemic lesions (lesion volume > 70 ml) and small ischemic lesions (lesion volume < 70 ml). In ICH, the objective is to identify patients who have a high risk for expansion and therefore are suited for early invasive treatment.

Research questions:

- How does a support vector machine based on quantitative DECT features compare to conventional, qualitative markers in predicting hematoma expansion in patients with intracerebral hemorrhage?
- To which extent can a deep convolutional neural network characterize the size and shape of the ischemic lesion on CT acquisitions in patients with acute ischemic stroke compared to DWI images?

Subquestions:

- Which DECT-based features are individually most accurate in predicting hematoma expansion in patients with intracerebral hemorrhage?
- How do DECT-based features rank in predicting hematoma expansion in a model consisting of multiple features?
- Which design choices for the architecture of a deep convolutional neural network encourage detection of relevant features for the detection of acute ischemic stroke and reduce noise?

Chapter 3

Early identification of expanding intracerebral hemorrhage on dual-energy CT using machine learning

3.1 Study goal

Hematoma expansion in ICH is individually associated with worse outcome and neurologic deterioration, and may benefit from early and aggressive medical or surgical intervention. There have been several clinical trials recently exploring medical and surgical treatment in ICH [18, 98, 22, 19]. Active blood pressure lowering, medical intervention and early surgical intervention did not improve the clinical outcomes or long-term benefits [99, 100, 18, 22]. Developing a method for early identification of patients at high risk for hematoma expansion who are most likely to benefit from early medical or surgical intervention can improve stratification of patients and thereby improve treatment efficacy [18]. In the past, several radiographic markers have been proposed to stratify patients for risk of hematoma expansion. The majority of described markers reflect baseline hematoma volume, contrast medium extravasation [47, 48, 49] and differences in hematoma morphology or the spatial configuration of intensities within a hematoma [40, 43, 41, 42, 44, 45, 46]. As of today, none of all these markers are able to show good predictive value of identifying patients who will likely develop hematoma expansion with both high sensitivity and specificity.

Dual-energy computed tomography (DECT) angiography enables separation of pixel intensities into hemorrhage, brain and contrast medium through three-material decomposition, utilizing the relative attribution of the photoelectric effect and Compton scattering to the energy-specific attenuation [101, 102]. This allows generation of separate virtual non-contrast (VNC) images and iodine-only maps (IOM), which can be used for quantification of contrast medium, see fig. 3.1. In this way, the IOM is able to provide unique, quantitative information about the distribution of contrast medium, while the VNC allows texture analysis of the hematoma. Combined with a semi-automatic delineation of the hematoma, it is possible to give a quick, standardized and quantitative assessment of the predictive value of several DECT-based features for hematoma expansion.

The aim of this study is to create a model of quantitative DECT-based features to predict hematoma expansion with higher sensitivity and specificity than current markers. Moreover, we assess the individual predictive value of different quantitative DECT-based features, their correlation and their relative rank of importance in predicting hematoma expansion.

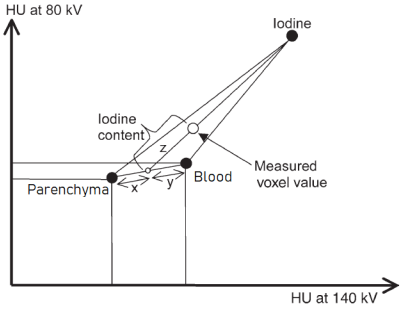

Figure 3.1: Three material decomposition decomposing voxel intensities into either blood or parenchyma, and attributing any deviations from the interpolation between blood and parenchyma to the presence of iodine [101]

3.2 Methods

3.2.1 Patient selection

Patients were included if they had intracerebral hemorrhage confirmed on CT angiography and follow-up CT. Exclusion criteria included: surgical intervention before follow-up imaging, placement of intraventricular drain before baseline scan, time between baseline and follow-up imaging of more than 48 hours and patients with exclusively intraventricular hemorrhage. The dataset was divided into a training set and test set, both obtained during different periods of time and are therefore independent from each other.

3.2.2 Image acquisition & hematoma classification

All DECT images were obtained using a Somatom Definition Force scanner (Siemens Healthcare, Forchheim, Germany) and post-processed using SyngoVia (Syngo Dual Energy Brain Hemorrhage, Siemens Healthcare) with the following scan protocol: tube A at 80 kVp, 499 mA, tube B at 140 kVp, 118 mA, with a collimation of 14 $\times$ 1.2 mm. Each image stack consisted of 80 kVp and 140 kVp series, both with a voxel resolution of 1 mm^3. The image stacks were post-processed to perform three-material decomposition using water and hemorrhage as base materials, attributing measured deviations from the linear combination of the attenuations of the two base materials to the presence of iodine, thereby reconstructing a VNC image and an IOM at a voxel resolution of 3 mm^3. The IOM in Hounsfield units were converted to iodine concentration images (in $\frac{mg}{ml}$) by fitting a linear regression to the calibration curve. Hematomas were classified as "expanders" if expansion was $> 3\ cm^3$ or $> 25\%$, and the classification was confirmed by two experienced neuroradiologists blinded to data analysis. For patients whose follow-up hematoma volume could not be calculated using our automatic segmentation algorithm because of the presence of a ventriculostomy catheter, the classification (expansion or stable) was determined by consensus between two neuroradiologists who independently and blindly reviewed the images to assess for hematoma expansion. For comparison, two neuroradiologists independently and blindly determined the presence or absence of the conventional spot sign in each case using previously described strict radiological criteria: a) one or more foci of contrast pooling within the ICH, b) with an attenuation $>$ 120 Hounsfield units, c) discontinuous from normal or abnormal vasculature adjacent to the ICH, and d) of any size and morphology [48].

3.2.3 Feature engineering

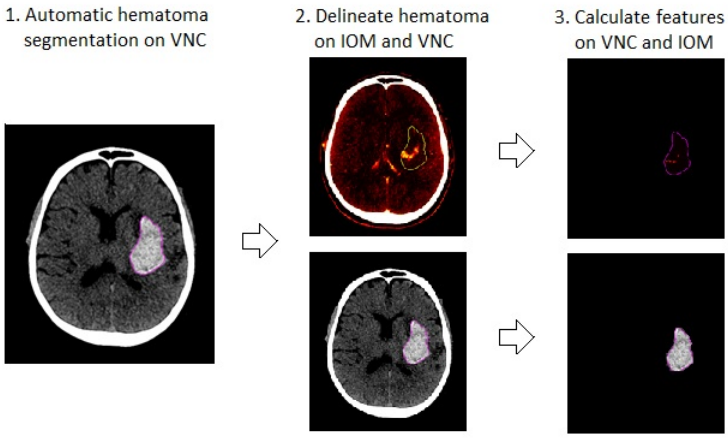

Figure 3.2: Visualization of the feature engineering process. The hematoma is delineated using a 3D continuous max-flow min-cut algorithm for multiple labels [103, 104, 105]. The delineation is superimposed on the IOM and VNC images, computing all features from the delineated hematoma. Iodine-based features in the hematoma calculated on the IOM are: iodine content in the hematoma, mean iodine content in the hematoma, iodine content in all spots, mean iodine content in all spots, iodine content in the brightest spot and the maximum voxel value in all spots. Features based on the hematoma texture on the VNC are: mean, variance, kurtosis, skewness, texture entropy, texture energy, texture dissimilarity, texture contrast and texture homogeneity.

Image analysis was performed in Matlab r2017b (The Mathworks inc., Natick, MA, USA). VNC images were semi-automatically segmented into skull, brain parenchyma and hematoma using a continuous max-flow algorithm to solve the 3D continuous-cut problem with multiple labels [103, 104, 105]. The boundary of the hematoma in the VNC image stack was used to delineate the hematoma on both IOM and VNC, see fig. 3.2. Based on the hematoma delineation on the VNC and IOM images, a variety of quantitative features were computed reflecting focal iodine extravasation, diffuse iodine extravasation and hematoma texture. Spots of iodine extravasation within the boundary of the hematoma were automatically identified in the IOM, using a quantile filtered mixture separation of a gamma distribution, representing spot sign pixel values, from a uniform distribution, representing brain parenchyma pixel values [106]. To prevent classifying noise or anatomic structures as spots, the connected components of pixels assigned to the gamma distribution need to fulfill size constraints in order to be classified as spots. The brightest spot was identified as the connected component within the identified spots with the highest mean pixel value. Iodine content in all spots, iodine content in the brightest spot and maximum voxel value in the spots were calculated to quantify the amount of focal iodine extravasation. Iodine content in the hematoma was calculated to quantify diffuse extravasation of iodine. To analyze the hematoma texture on VNC, mean, variance, kurtosis, skewness, texture entropy, texture energy, texture dissimilarity, texture contrast and texture homogeneity were computed. Mean, variance, kurtosis and skewness are computed based on the histogram of the hematoma in 3 dimensions on the VNC. Entropy, energy, dissimilarity, contrast and homogeneity are based on a 2D gray-level co-occurrence matrix [107].

3.2.4 Feature selection

Statistical analyses were performed using R language for statistical computing. Features were assessed for normality using the Shapiro-Wilk test. Since the majority of features is non-normally distributed, for the sake of comparability all feature values are presented as median with their corresponding interquartile range (IQR). To create a prediction model to classify hematomas as expanders or stable, a classifier has to be constructed as a function of a subset of the available features. Therefore, the best performing subset of features has to be determined according to some criterion. For all individual features, their accuracy in classifying hematomas as expanders or non-expanders was computed using the area under the curve (AUC) of the receiver-operator characteristic (ROC). A feature was only retained if the 95% confidence interval of the AUC did not cross the random chance rate, equal to 0.5. Between all individual features a correlation coefficient was computed, eliminating the individually worst performing feature (measured using by their respective AUC) from a pair of predictors whose correlation coefficient was above 0.95 or below -0.95. On the remaining individually predictive features, a linear regression-based forward stepwise parameter selection approach was used, optimizing criteria such as the optimizing the coefficient of determination (R^2), adjusted coefficient of determination (adjusted R^2) and root mean square of the error between the predicted outcome and the supervision [108]. Starting with an empty model, at each iteration the forward parameter selection selected the parameter with the lowest p-value for the F-test between the existing model and the model plus that parameter.

3.2.5 Model creation

Using the combination of features derived from forward parameter selection, we defined a classification model based on a support vector machine learning algorithm for classifying hematomas as expanders or stable. The classification boundary was constructed automatically, implicitly maximizing the margin between the different classes in higher-dimensional feature space using a radial basis function kernel K, defining the dot product of two feature vectors in feature space as:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma|\mathbf{x}_i - \mathbf{x}_j|^2). \tag{3.1}$$

With $\mathbf{x}_i$ and $\mathbf{x}_j$ being the input feature vectors, and γ a constant defining the influence of a single data sample on the model. The classification boundary was constructed on the training set. The SVM model was then validated on an independent data set. The training set and test set were acquired during two different periods of time, and therefore independent. A confusion matrix was used to evaluate the performance of the classification model on the training set and on the test set and to compare the performance to the presence of conventional spot sign, reporting sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and accuracy.

3.3 Results

3.3.1 Patient population

Between October 2014 and February 2018, a total of 176 patients who had ICH were referred to DECT angiography for assessment. Of these patients, 40 had no follow-up CT within 48 hours, 4 had craniotomy or hematoma evacuation before follow-up, 20 had exclusively intraventricular hemorrhage, 6 had an extraventricular drain at the initial scan and 1 had inadequate image quality. The 105 remaining patients had 114 hematomas (1 patient had 6 hematomas, 1 patient had 3 hematomas and 2 patients had 2 hematomas each) of which 28 hematomas expanded while 86 did not (table 3.1). The dataset was divided in a training set of 41 patients (45 hematomas), acquired between October 2014 and January 2017, and an independently obtained test set 64

Table 3.1: Distribution of predictors for expanding and stable hematomas

	Stable (n = 86)	Expanding (n = 28)	P-value
Iodine quantification on IOM			
Total iodine in hematoma (mg)	3.39 (0.91 – 7.32)	20.61 (8.35 – 30.16)	<0.01
Total iodine in focal spots (mg)			
All spots	0.0034 (0 – 0.069)	0.49 (0.12 – 1.05)	<0.01
Brightest spot	0.0024 (0 – 0.025)	0.044 (0.022 – 0.21)	<0.01
Mean iodine in hematoma (mg/ml)	0.44 (0.38 – 0.64)	0.48 (0.41 – 0.61)	0.58
Mean iodine in all spots (mg/ml)	1.45 (0 – 2.16)	2.34 (1.86 – 3.04)	<0.01
Max pixel value in focal spots (mg/ml)	4.93 (2.85 – 7.74)	1.48 (0 – 3.56)	<0.01
Hematoma analysis on VNC			
Mean	56.54 (54.56 - 59.54)	56.36 (51.65 - 60.22)	0.36
Variance	53.09 (41.88 - 64.53)	50.14 (40.80 - 61.46)	0.42
Kurtosis	2.81 (2.56 – 3.09)	2.95 (2.59 – 3.23)	0.56
Skewness	-0.32 (-0.56 – -0.086)	-0.35 (-0.59 – -0.17)	0.75
Texture entropy	5.57(5.42 – 5.72)	5.49 (5.32 – 5.67)	0.25
Texture energy	0.0054 (0.0046 - 0.0063)	0.0058 (0.0047 - 0.0072)	0.30
Texture dissimilarity	1.70 (1.61 - 1.86)	1.73 (1.57 - 1.96)	0.50
Texture contrast	0.024 (0.014 – 0.035)	0.052 (0.033 – 0.071)	0.38
Texture homogeneity	1.00 (1.00 – 1.00)	1.00 (1.00 – 1.00)	0.32

patients (69 hematomas), acquired between February 2017 and February 2018. The 105 included patients had a median time before follow up of 7 hours (interquartile range: [5 – 16]).

3.3.2 Feature selection

We designed the following features based on the VNC and IOM of the DECT angiography: iodine content in the hematoma (I_h), iodine content in the brightest spot (I_{bs}), iodine content in all spots (I_s), mean iodine content per voxel in all spots (mean I_s), max iodine content per voxel in all spots (max I_s), mean, variance, kurtosis, skewness, texture entropy, texture energy, texture dissimilarity, texture contrast, texture homogeneity.

Individual predictive value of predictors

The median and interquartile range of every predictor for both expanders and non-expanders are presented in table 3.1. The medians were significantly different ($p < 0.05$) between the stable and expanding hematomas for the following predictors: I_h, I_s, I_{bs}, mean I_s and max I_s. For each predictor, the area under curve (AUC) of the receiver-operator characteristic (ROC) curve is visualized in fig. 3.3. The features considered for modeling of hematoma expansion are the features whose confidence interval did not cross the random chance rate (AUC = 0.5), which were in order of decreasing AUC: I_s, I_{bs}, I_h, max I_s and mean I_s.

Feature correlation

In order to exclude predictors which do not contain exclusive predictive information, the correlation coefficient was computed among all the predictive features, see fig 3.4. The following

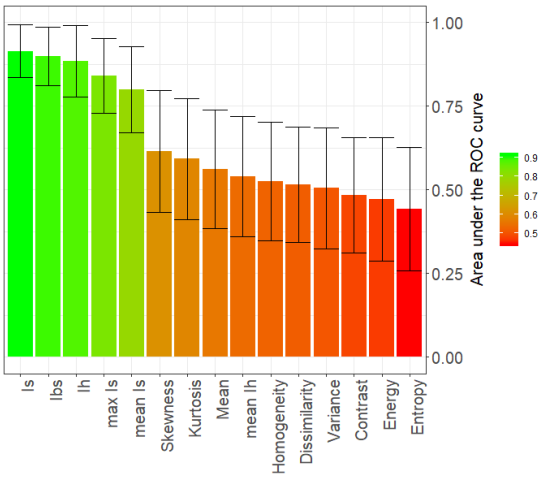

Figure 3.3: AUC of the ROC for all individual features. Only features whose 95% confidence interval does not cross the random chance rate of AUC = 0.5 are considered for expansion modeling.

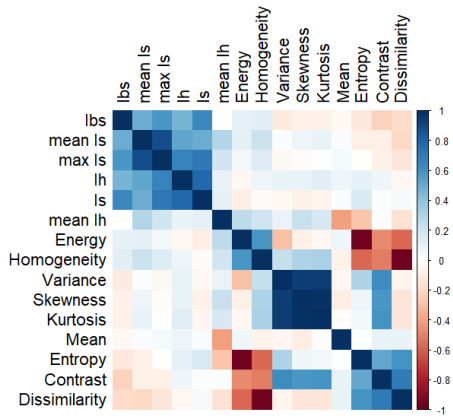

Figure 3.4: Visualization of the correlation coefficient between features. Strongly correlated features turn dark red or dark blue. If the correlation coefficient between two features is $r > 0.95 \cup r < -0.95$, the individually least predictive feature is dropped for hematoma modeling.

features had a too strong correlation to be considered individually predictive: texture homogeneity & texture dissimilarity ($r = -0.96$), texture energy & texture entropy ($r = -0.96$), skewness & kurtosis ($r = 0.98$), variance & kurtosis ($r = 0.95$), variance & skewness ($r = 0.95$). The excluded features based on correlation coefficient were: texture dissimilarity, texture entropy, kurtosis and variance.

Optimal combination of features for prediction model

On the remaining individually predictive features, a linear regression-based stepwise parameter selection approach was used without a limitation on maximum p-value for the F-test. To determine the optimal number and combination of features, the coefficient of determination (R^2), adjusted coefficient of determination (adjusted R^2) and root mean square of the error between the predicted values and the actual values (RMSE) were computed for every number of features ranging from 1 to all candidate features. Starting with an empty prediction model, at each iteration the feature was selected which explained the largest portion of variance in outcome in addition to the existing model, having the lowest p-value for the F-test. This process yielded a ranking of relative importance of features in explaining outcome variance, taking into account interactions among the features. This ranking in order from most to least important feature was: I_h, I_{bs}, I_s, mean I_s, max I_s, hematoma contrast, hematoma homogeneity and hematoma entropy. Fig. 3.5 shows that R^2, the proportion of explained variance in outcome, kept increasing with each feature addition to the model. Adjusted R^2 corrects for the modeling of random chance by additional features, showing only the proportion of explained variance in outcome larger than the proportion which can be modeled by random chance. Adjusted R^2 was maximal at a model of 3 features: I_h, I_{bs} and I_s. The RMSE between the predicted values and the actual values was minimal at the combination of the same 3 features. After recomputation of the forward parameter selection, selecting only those parameters whose p-value is smaller than 0.05, the only parameters included in the model were: I_h and I_{bs}.

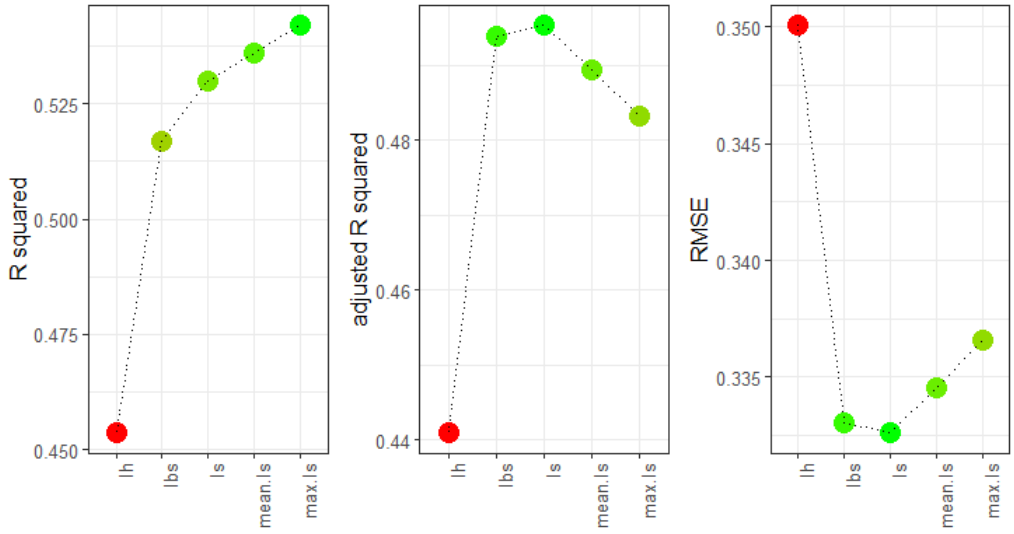

Figure 3.5: Metrics of model performance as a function of the amount of features included in the model. The metrics for measuring model performance were: the coefficient of determination (R^2), adjusted coefficient of determination (adjusted R^2) and the root mean square of the error (RMSE). The model showed optimal performance at 3 features, I_h, I_{bs} and I_s, minimizing adjusted R^2 and RMSE.

3.3.3 Model performance

The constructed support vector machine classification model was based on I_h and I_{bs}, and is visualized in fig. 3.6. The classification boundary separated the training set of 45 hematomas in stable and expanding hematomas, correctly identifying 30/31 stable hematomas (specificity of 97%) and 10/14 expanding hematomas (sensitivity of 71%), see table 3.2. Using the support vector machine classification model tuned on the training set, the model separated the independently acquired test set of 69 hematomas into stable and expanding hematomas, correctly identifying 50/55 stable hematomas (sensitivity of 91%) and 7/14 expanding hematomas (sensitivity of 50%), see table 3.2.

Conventional spot sign performance

The performance of the conventional spot sign was also evaluated on the training and test set, see table 3.3. In total, 12 out of 88 stable hematomas were marked with a spot sign, and 16 out of 28 expanding hematomas were marked with a spot sign. The conventional spot sign separated the training set of 45 hematomas in stable and expanding hematomas, correctly identifying 8/14 expanding hematomas (sensitivity of 57%) and 28/31 stable hematomas (specificity of 90%). On the test set, the conventional spot sign identified 8/14 expanding hematomas correctly (sensitivity of 57%) and 47/56 stable hematomas correctly (specificity of 83%).

3.4 Discussion

In this study, we designed, identified, and evaluated quantitative DECT-based features of spontaneous intracerebral hemorrhage for prediction of hematoma expansion. We evaluated the perfor-

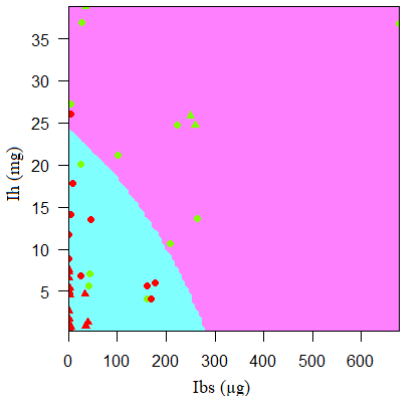

Figure 3.6: Visualization of employed classification model. Predicted expansion is shown in purple and predicted stability is shown in blue, actual expansion is shown in green and actual stability is shown in red. Data points functioning as support vectors and therefore have an influence on the position of the classification boundary are shown as a circle, other data points are shown as a triangle.

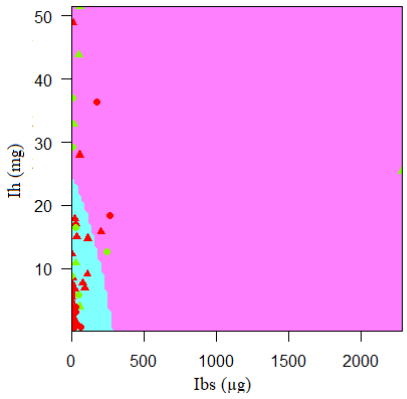

Figure 3.7: Visualization of classification model on test data. Predicted expansion is shown in purple and predicted stability is shown in blue, actual expansion is shown in green and actual stability is shown in red. The classification boundary is constructed during training and is equal to the classification boundary in fig. 3.6

mance of our quantitative model by comparing it to the conventional spot sign. Furthermore, we assessed the individual and cumulative predictive value of our quantitative DECT-based features.

Our quantitative approach employing a support vector machine model had a superior sensitivity and specificity than the conventional spot sign on the train set as classified by our neuroradiologists and what is reported in literature (50% - 60% sensitivity and 80% - 90% specificity), a superior specificity on the test set than the conventional the spot sign in our cohort and in literature, but a lower sensitivity on the test set than the conventional spot sign in our cohort and comparable to the sensitivity reported in literature [109, 48, 49, 110, 47, 36]. The performance of the conventional spot sign in our cohort is comparable to the performance reported in this literature. Assuming that the set of feature vectors and corresponding output, $(\mathbf{x}, y)$, are independent and identically distributed and drawn from a fixed but unknown probability distribution, $\mathrm{P}(\mathbf{x}, y)$, the decrease in performance of our model is caused by an insufficient amount of training or test data. An insufficient amount of data implicates that the set of feature vectors and outcomes $(\mathbf{x}, y)$ is not an accurate representation of the actual probability distribution in the population $\mathrm{P}(\mathbf{x}, y)$. An increase in the amount of data reduces the performance gap between train and test set, improving model generalization [97]. A different explanation for the performance drop in sensitivity between train and test set is that the data samples are not independent or identically distributed. Varying settings among different CT acquisitions, such as contrast injection rate and reconstruction kernel, influence the quantification of iodine and therefore changes $\mathrm{P}(\mathbf{x}, y)$ between different acquisitions. Moreover, the evaluation of prediction probability is based on single hematomas, and there are multiple cases where a patient has multiple hematomas which are considered independent data samples while they share an underlying physiology and therefore do not meet the criterion for independent features. However, this occurs in both the train as the test set, and would be unlikely to cause the observed drop in model performance.

We designed and evaluated 15 quantitative dual-energy CT-based features, and included 2 of those features in our classification model, the iodine content in the hematoma and the iodine content in the brightest spot. The cumulative model performance of the the candidate features was computed by forward stepwise selection, implicitly defining a rank of feature importance when the interactions with other, already included features were considered. This ranking was: iodine content in the hematoma, iodine content in the brightest spot sign, iodine content in all spots, maximum voxel value in all spots and mean voxel value in all spots, see fig. 3.5. The features included in the model were the only features that could be included in a linear model with p-value < 0.05. However, the best performing model according to adjusted R^2 and RMSE was with the addition of a third feature, the iodine content in all spots. Although this third variable had explained variance in the outcome variable superior than random chance, demonstrated by the increase in adjusted R^2, this explanation of outcome variance apparently was not sufficient to effectuate a p-value < 0.05. Addition of other features decreased the generalizable performance of the model, indicated by an increase of R^2, decrease of adjusted R^2 and decrease of RMSE as shown in 3.5. The most important predictor in our model was the iodine content in the hematoma, indicating that diffuse iodine extravasation in the hematoma is an important marker for hematoma expansion, in addition to the currently well known focal spot sign [29, 36, 48]. Since hematoma expansion can be due to either focal or diffuse extravasation of blood [111], the quantification of iodine content in the hematoma is a new, DECT-based method to model diffuse iodine extravasation. The possibility to more accurately capture the physiology behind hematoma expansion could improve identification of patients at risk by recognition of diffuse bleeding in addition to the focal bleeding of the conventional spot sign. The second most important feature in our model was the iodine content in the brightest spot, reflecting focal iodine extravasations. The iodine content in the brightest spot was identified to be more important for modeling of hematoma expansion than the total content of iodine in all spots, contradicting the current practice of characterizing spot signs by their number, size and intensity into a composite score [109]. Our results indicate that the intensity of the brightest spot may be a better indicator for hematoma expansion than number and size of the spots. In this analysis, the stepwise forward parameter selection used a linear regression to compute the p-values, possibly over- or underestimating the importance of features that require different modeling than a linear model. The p-values computed during stepwise forward parameter selection are a measure of difference in outcome variance. However, since our output can only assume two states, expansion or stable, outcome variance might not be an adequate metric to evaluate feature importance.

The classification accuracy of the 15 quantitative DECT-based features was also computed individually for each feature and expressed in terms of area under the receiver-operator charac-

Table 3.2: Model performance

	Predicted expansion	Predicted stable	
	Training set		
Expansion	TP = 10	FN = 4	Sensitivity = 71%
Stable	FP = 1	TN = 30	Specificity = 97%
	PPV = 91%	NPV = 88%	Accuracy = 89%
	Test set		
Expansion	TP = 7	FN = 7	Sensitivity = 50%
Stable	FP = 5	TN = 50	Specificity = 91%
	PPV = 58%	NPV = 88%	Accuracy = 83%

Table 3.3: Conventional spot sign performance

	Spot sign	No spot sign	
	Training set		
Expansion	TP = 8	FN = 6	Sensitivity = 57%
Stable	FP = 3	TN = 28	Specificity = 90%
	PPV = 73%	NPV = 82%	Accuracy = 80%
	Test set		
Expansion	TP = 8	FN = 6	Sensitivity = 57%
Stable	FP = 9	TN = 48	Specificity = 84%
	PPV = 47%	NPV = 89%	Accuracy = 79%

teristic. The rank for the individual accuracy of the features in order of decreasing AUC was: iodine content in all spots, iodine content in the brightest spot, iodine content in the hematoma, maximum voxel value in all spots, mean voxel value of all spots, skewness, kurtosis, mean, mean iodine content in the hematoma, homogeneity, dissimilarity, variance, contrast, energy and entropy. A clear distinction in the individual accuracies was that the 5 features with highest AUC were based on quantifications of iodine on the IOM, while all 9 features based on hematoma texture on the VNC had a lower AUC and had a 95% confidence interval that crossed random chance rate. This indicates that features based on iodine quantification capture physiology underlying hematoma expansion better than hematoma-texture-based features. The quantitative features based on hematoma texture are not able to capture the qualitative-texture-based signs reported in literature, such as [111, 112, 113]. Please note that a different study did find a predictive value for similar quantitative hematoma-texture-based features in predicting hematoma expansion [114]

Our study has multiple limitations. First, we only obtained data from 105 patients, inherently limiting the ability of our model to generalize to new data. Second, the images were acquired on only 1 scanner in 1 hospital, making the derived model sensitive to different scanners, hospital settings, acquisition methods and scan protocols. Third, we have treated and counted multiple, non-contiguous regions of intracerebral hemorrhage in any patient as separate, independent occurrences of hemorrhage. While the DECT-based features for these hematomas in one patient can be computed independently, they are not independent as they are subject to a common underlying physiology. Fourth, acquisitions in our cohort had a non-trivial variability in scan settings and acquisition timing. Different scans were reconstructed with different kernels, with different kernels highlighting high-frequent or low-frequent spatial information and subsequently affecting the identification and characterization of spot signs and the nature of hematoma texture. On top of that, the variability in the delay time from initial CT angiography acquisition to the delayed CT angiography acquisition influenced the appearance of iodine in the CT scan. Since timing of delayed imaging can impact the identification of spot signs, our features were probably not immune to this variability in delay time [48, 115].

Future research should focus on evaluating quantitative DECT-based features with independent and identically distributed data samples where the data samples in the train and test set approximate the underlying probability distribution. Therefore, several changes have to be made with respect to this study. To start, the amount of data samples has to be increased to better approximate the probability distribution from which they are drawn. Especially since expanding hematomas are relatively scarce compared to stable hematomas, the current amount of data is not enough to construct a generalizable model. Moreover, to make sure that data samples are

drawn from an identical distribution, acquisitions should be standardized with an identical protocol and reconstruction settings. To enforce that data samples are independent, only patients with 1 hematoma should be included in the study. If the added value of quantitative DECT-based features is proven in such a controlled environment, protocol constraints can be relaxed to evaluate the value of these features in clinical practice.

3.5 Conclusion

We have demonstrated that a support vector machine, constructed from quantitative DECT-based features, can identify expanding hematomas better than the conventional spot sign on the train set and similar to the conventional spot sign on the test set. The support vector machine was constructed using only 2 features: iodine content in the hematoma and iodine content in the brightest spot. All the features with individual classification accuracy whose confidence interval did not cross random chance rate were based on iodine quantifications, indicating that hematoma texture is either less useful for predicting hematoma expansion than iodine-based features or that our texture-based features did not capture the actual hematoma texture correctly.

This research shows the potential of quantitative DECT-based features in the prediction of hematoma expansion, which can be used to stratify patients with intracerebral hemorrhage for invasive treatment to increase therapeutic efficacy.

Chapter 4

Identification of acute ischemic stroke on CT angiography using deep learning

4.1 Study goal

Whether a patient suffering from AIS is eligible for EVT is based on the extent of infarction observed on DWI in combination with the patient's clinical presentation. Although presumably superior in supporting clinical decision making compared to other modalities, DWI is widely constrained due to workflow and availability issues. The search for modalities able to provide similar information about the extent of the parenchymal lesion as DWI while using less resources sparked interest in further research into CT angiography to quantify the patient's collateral function in AIS. It has previously been shown that better collateral function in AIS is associated with small ischemic core, large clinical penumbra and slower progression of infarcted tissue [85, 86]. Since collateral function is related to ischemic core and penumbra size, a comprehensive assessment of the collateral function in AIS patients may aid in patient selection for EVT without the necessity to utilize CT perfusion or MRI-based acquisitions [85, 86]. However, currently only qualitative methods describing grading scales of the affected vascular territory exist for evaluating collateral function [90, 85]. A more quantitative approach of assessing iodine distribution in CT angiography images and delayed phase CT angiography images could open up possibilities to rethink patient selection for EVT in AIS.

The non-contrast CT, CT angiography and delayed-phase CT angiography contain information about the blood supply in non-affected brain tissue and ischemic brain tissue to potentially aid in clinical decision making for the triage of AIS patients. The primary challenge is to derive a model based on relevant and generalizable features able to predict whether a patient is eligible for EVT. In medical image analysis, deep learning methods have proven to be highly effective in classification, localization and segmentation when the amount of available labeled training data is large enough to learn complex patterns [116, 117]. The aim of this study is to optimize a deep convolutional neural network to determine the presence, extent and location of acute ischemic stroke based on CT and CT angiography.

4.2 Methods

4.2.1 Patient selection

Patients were included if they were referred to CT and subsequently to MRI for the evaluation of acute stroke symptoms. Exclusion criteria included: CT or MRI acquisitions that did not image the entire head and inadequate delay times between CT angiography and delayed-phase CT angiography ($<$ 15 seconds or $>$ 60 seconds). Patients were divided in either train set or

test based on the presence of a large hematoma (infarct volume on follow up DWI > 30 ml). 75% of patients with a large hematoma and 75% of patients with a small or no hematoma were included in the train set, 25% of patients with a large hematoma and 25% of patients with a small hematoma were included in the test set.

4.2.2 Image acquisition & image analysis

CT images were obtained using a both a DECT scanner and a regular CT scanner. Since data was obtained retrospectively, no clear-cut protocol defining voxel size, reconstruction kernel and tube voltage was applicable to all scans. The default imaging protocol for the assessment of acute ischemic stroke on this scanner consisted of a non-contrast CT, bolus-tracked CT angiography, and a delayed-phase CT angiography with a 20-30 seconds time delay. DWI images were obtained using multiple MR scanners, with varying protocols for acquisition of DWI images. Rigid registration of non-contrast CT, CT angiography, delayed CT angiography and diffusion weighted MRI to a standard space was performed using SPM12 (Wellcome Trust Centre for Neuroimaging, University College London, London, United Kingdom). In standard space, each image stack had 28 slices, with a slice thickness of 5 mm, of 512 $\times$ 512 voxels with an axial resolution of 1 $\times$ 1 mm. Segmentation of DWI images into ischemic lesion masks for supervised learning was performed semi-automatically in Matlab using a custom algorithm and was supervised by a neuroradiologist.

4.2.3 Deep learning

All deep learning was performed using a custom algorithm in Keras (F. Chollet, https://keras.io) and TensorFlow (TensorFlow, https://www.tensorflow.org) running in Python 3.5 (Python Software Foundation, https://www.python.org/).

Convolutional neural network

To learn the identification of AIS on CT images in a supervised fashion, the CT input was passed through a convolutional neural network (CNN). The CNN is a 16 layer encoder-decoder network, denoted as $\mathbf{X}$. The layers are indexed from 0 to L, with layer 0 being the input layer and layer L the output layer. The state of layer i is denoted as $\mathbf{x}_i$. The entire network is constructed by concatenation of layers $\mathbf{x}_i$ for $i = 0, \cdots, L$. The basic operations of the employed CNN are the convolution and deconvolution, which both map a feature map i to a subsequent feature map j using a Rectified Linear Unit (ReLU) as activation function:

$$\mathbf{x}_j = max(0,\ \mathbf{w}_{ij} \circledast \mathbf{x}_i + \mathbf{b}_j). \tag{4.1}$$

Where $\mathbf{W}$ represents the weights matrix between each feature map i and feature map j, and $\mathbf{b}_j$ is the bias of feature map j. However, since each convolution is followed by batch normalization, $\mathbf{b}_j$ will be canceled out by mean subtraction and can thus be ignored [118]. Leading to the following expression for a convolution followed by batch normalization (BN):

$$\mathbf{x}_j = max(BN(0,\ \mathbf{w}_{ij} \circledast \mathbf{x}_i)). \tag{4.2}$$

Where batch normalization serves the purpose to standardize the input to each layer in the network to zero mean and unit variance, accelerating network training by allowing higher learning rates and be less careful about parameter initialization [118]. Consider the input $\mathbf{x}$ passed through the CNN of L layers, each implementing a non-linear transformation H_l at layer l. Layer l maps a given layer input $\mathbf{x}_l$ to a representation $\mathbf{x}_{l+1}$ which serves as the input for the subsequent layer. H_l can be a composite function of operations such as transposed convolutions, convolutions,

batch normalizations and ReLUs. If we denote the output of layer l as $\mathbf{x}_l$, considering we have a short skip connection inside the layer l and a long skip connections between each layer l and $L-l$, each encoding layer l can be expressed as:

$$\mathbf{x}_{l+1} = [H_l(\mathbf{x}_l),\ \mathbf{x}_l]. \tag{4.3}$$

Long skip connections between each layer l and $L-l$ serve to pass on low-level features and eliminate gradient singularities that slow down training but are inherent in nodes in a deep networks, while short skip connections inside each layer enhance optimization of deeper networks by allowing uninterrupted gradient flow [119, 120]. Because each decoding layer receives a skip connection from an encoding layer, for the decoding layer $L-l$ follows:

$$\mathbf{x}_{L-l+1} = [H_{L-l}([\mathbf{x}_{L-l}, \mathbf{x}_l]),\ \mathbf{x}_{L-l},\ \mathbf{x}_l]. \tag{4.4}$$

The encoding part of the network performs convolutions and max pooling operations resulting in a learned downsampling of the feature maps, while the decoding part of the network performs transposed convolutions and convolutions, serving to perform a learnable upsampling. The currently implemented model is visualized in fig. 4.1. The topography of both the encoding and

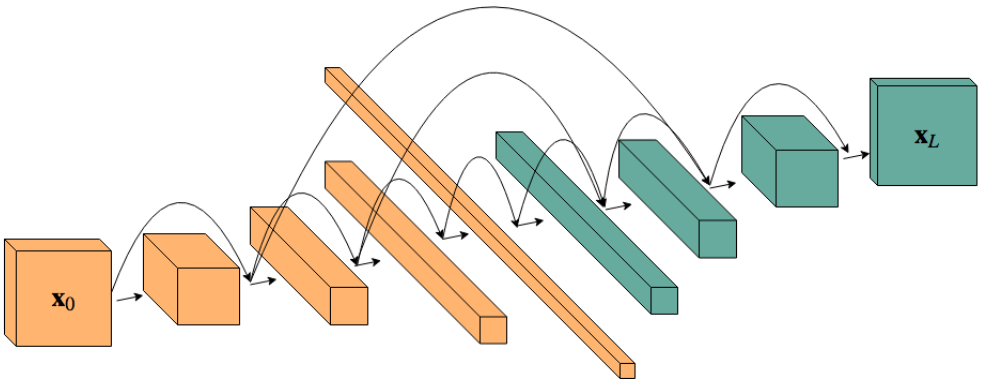

Figure 4.1: Schematic representation of the convolutional neural network. The orange layers are the encoding part of the network performing a learnable downsampling, the green layers are the decoding part performing a learnable upsampling. Black arrows are feature map concatenations either within a layer, known as short skip connections, or between layers l and $L-l$, known as long skip connections.

decoding layers are visualized in fig. 4.2. The network input is a $u \times v \times 3$ volume, with the individual channels consisting of non-contrast CT, CT angiography and delayed-phase CT angiography. The supervision is a $u \times v$ mask of infarcted lesion, denoted as $\mathbf{R}$, with supervision pixel values $\mathbf{r}_{u,v}$. The output is a $u \times v$ pixel-wise probability distribution of a pixel being infarcted $\mathbf{P}$, with pixel values $\mathbf{p}_{u,v}$. $\mathbf{P}$ is optimized to match $\mathbf{R}$ by minimizing the generalized dice loss, which has shown to be most robust against class imbalance in the data [121]. The generalized dice loss (GDL) is:

$$GDL = 1 - 2\frac{\sum_{l=1}^{2} w_l \sum_u \sum_v \mathbf{r}_{l,u,v}\mathbf{p}_{l,u,v}}{\sum_{l=1}^{2} w_l \sum_u \sum_v \mathbf{r}_{l,u,v} + \mathbf{p}_{l,u,v}}. \tag{4.5}$$

Where each label l has a label weighting to correct for class imbalance, w_l, inverse to its relative volume as:

$$w_l = (\sum_u \sum_v \mathbf{r}_{l,u,v})^{-2}. \tag{4.6}$$

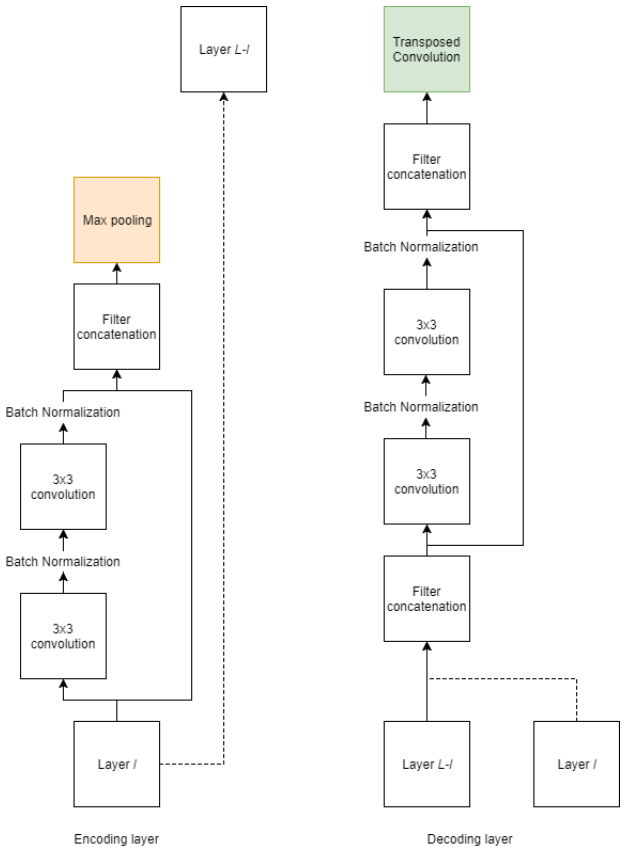

Figure 4.2: Topography of the encoding and decoding layers. Encoding layer l has a short skip connection inside the layer and sends a long skip connection to layer $L-l$, learnable downsampling is performed during the max pooling operation. Decoding layer $L-l$ has a short skip connection inside the layer and receives a long skip connection from layer l, learnable upsampling is performed during the transposed convolution.

Implementing the Dice score in a two-class problem with SGD, as in our case, the gradient with respect to a pixel having probability of infarction p_i becomes:

$$\frac{\delta GDL}{\delta p_i} = -2\frac{(w_1, w_2)^2)(\sum_u \sum_v \mathbf{p}_{u,v}\mathbf{r}_{u,v} - r_i \sum_u \sum_v (\mathbf{p}_{u,v} + \mathbf{r}_{u,v})) + uvw_2(w_1 + w_2)(1 - 2r_i)}{((w_1 - w_2)\sum_u \sum_v (\mathbf{p}_{u,v} + \mathbf{r}_{u,v}) + 2uvw_2)^2}, \tag{4.7}$$

as shown in [121]. The gradient with respect to each preceding input layer can be subsequently computed using the chain rule.

4.2.4 Hyperparameters

The employed CNN was trained during 150 forward and backward passes through the network of all training samples, called epochs. Each forward and backward pass is computed over a number of training samples, called the mini-batch size, which was set to 8. The employed optimizer for weight adjustment was a stochastic gradient descent algorithm with momentum, using a learning rate scheduler to decrease the learning rate gradually as a function of the completed amount of

epochs with a starting learning of 0.1. The network depth of 16 layers and the amount of hidden units were not varied during training.

4.3 Results

4.3.1 Patient population

Between February 2017 and February 2018, a total of 101 patients were assessed for AIS using both CTA and MR. Of these patients, 16 patients had a delayed-phase CT angiography acquisition that did not contain the entire head, 11 patients had delay times appropriate for evaluation of intracerebral hemorrhage but not for AIS evaluation, 2 patients had an MR acquisition without DWI and 1 patient was excluded due to poor image quality. The remaining 71 patients were divided in either train set or test. The train set was composed of 53 patients, of which 5 had a large hematoma, and the test set was composed of 18 patients, of which 2 had a large hematoma.

4.3.2 Identification of acute ischemic stroke

The CNN was trained on 53 patients, minimizing the Dice loss by gradually matching the continuous function network output to the ischemic lesion mask supervision. The Dice loss gradually decreased from almost 1 at epoch 1 to 0.32 at epoch 150, see fig. 4.3. This Dice coefficient after

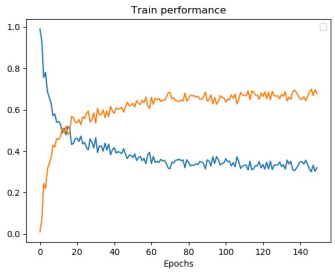

Figure 4.3: Visualization of the Dice loss (orange line) and the Dice coefficient (blue line) as a function of the amount of completed epochs during training.

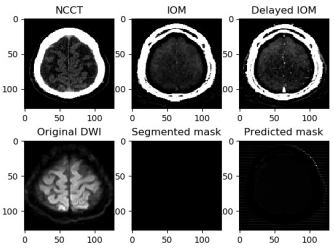

Figure 4.4: Example of the predicted model output on training data without large stroke volume.

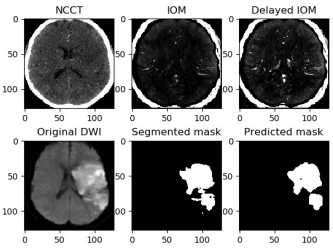

Figure 4.5: Example of the predicted model output on training data with a large stroke volume.

training loss means that there was a 68% union between the ischemic lesion on the supervision and the generated output. The results of the generated output compared to the supervision are presented in fig. 4.4 and fig. 4.5. After training, the deep learning model was validated on new, unseen data in the test set. However, the Dice loss on the validation set remained close to 1, as can be seen in fig. 4.6, with the results of the generated output compared to the supervision on the test set are presented in fig. 4.7 and fig. 4.8.

4.4 Discussion

In this study, we designed and evaluated a deep convolutional neural network aiming to characterize the size and shape of ischemic lesions in patients with acute ischemic stroke based on CT angiography acquisitions. The CNN was designed to detect abnormalities in anatomy and

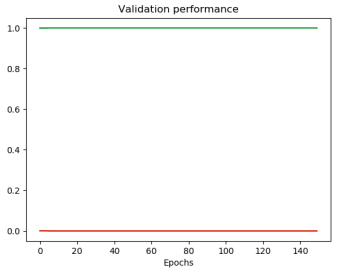

Figure 4.6: Visualization of the Dice loss (green line) and the Dice coefficient (red line) of the validation set as a function of the amount of completed epochs during training.

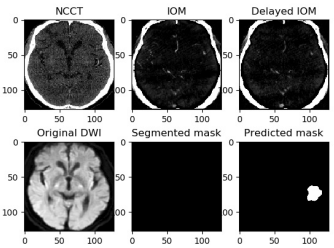

Figure 4.7: Example of the predicted model output on test data without large stroke volume.

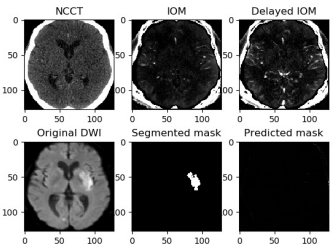

Figure 4.8: Example of the predicted model output on test data with a large stroke volume.

physiology in different CT acquisitions. We evaluated the performance of our CNN by computing a Dice loss for the train set and validation set as a function of the amount of epochs during training, as well as computing a Dice loss for the test set after training.

While our CNN showed promising performance in the identification of AIS on CT on the train set (Dice loss < 0.4), this performance could not be reproduced on the test set (Dice loss ≈ 1). The increasingly larger difference between Dice loss on the train set and Dice loss on the validation set during training indicates that the CNN overfitted on the training data. The expected generalization error, the difference in expected train error and expected test error, is shown to be proportional to $\frac{h}{P}$, where P is the number of training samples and h is a measure of complexity, or effective capacity, of the model [97]. The most probable cause of overfitting in this case is an insufficient amount of data presented during training. Although the amount of available data samples during training equals 53 patients $\times$ 28 slices = 1484 data samples, the large variety in slices from different, parallel parts of the brain in contrast to the subtle physiological changes in ischemic territory in combination with the low proportion of slices containing ischemic lesions make the problem of ischemic lesion characterization too complex to solve in a generalizable fashion with this amount of data. It is important to note that the effective capacity of CNNs has been shown to be sufficient for memorizing entire datasets of random noise [122]. The concept of regularization is to restrict the possible set of functions to a subset of less complex functions, reducing effective capacity and thereby improving generalization. The role of explicit generalization, such as weight decay, dropout, data augmentation and L1/L2 regularization, and implicit generalization, such as batch normalization and early stopping, is an active topic of discussion in deep learning [122]. In the employed CNN we incorporated both explicit generalization, weight decay, as implicit generalization, both batch normalization and early stopping, but it did not suffice to prevent a large generalization error. The implementation of data augmentation, synthetically generating 'new' data samples by performing affine transformations on existing data samples and their supervision, would increase the amount of data samples but also the complexity of the task for the CNN, since the data would not be registered to the same standard space. Implementing more regularization techniques to make the model more generalizable, such as dropout or L1/L2 regularization can help to reduce generalization error [123]. However, the generalization error appears to be too large to be solved with only regularization. A different approach would be to reduce the task complexity by delivering a more structured input to the network, such as only 2-dimensional slices of the same anatomical region for one CNN or only

3-dimensional volumes of the entire brain.

This study has several limitations. First and foremost, the amount of training and test data is too small for achieving a small generalization error. Second, the input data varies in the way the CTs are acquired. The reconstruction kernel, contrast injection rate, delay time between CT angiography and delayed-phase CT angiography and acquisition settings vary among data samples. Third, there is a variation in the etiology of ischemic lesions and the time since onset of stroke symptoms in our dataset, both causing a variation of input data.

Future research should focus on the trade-off of between generalization and network complexity for complex deep learning problems. With respect to this study, future research would benefit from improving standardization of input data and above all gather more training and test data to try and detect ischemic lesions based on CT acquisitions.

4.5 Conclusion

We have not been able to characterize the shape and size of ischemic lesions in a generalizable fashion. The amount of training data available was not sufficient to reach a small generalization error given the complexity of the characterization task and the complexity of the deep CNN. Future work should focus on the inclusion of more patients, standardization of input data in both acquisition protocol and stroke etiology and the adaptation of the CNN to optimize between network complexity and the ability to generalize to new data.

Chapter 5

Conclusion

In this study we attempted to develop CT-based machine learning models to identify subgroups of either patients at risk for hematoma expansion in intracerebral hemorrhage or patients with a large ischemic lesion in acute ischemic stroke. Identification of these subgroups of patients can potentially increase treatment efficacy in both hemorrhagic and ischemic stroke.

In order to identify expanding hematomas in patients with intracerebral hemorrhage, we employed an SVM-based machine learning model using quantitative DECT-based features. The model showed superior performance to the conventional spot sign on the train set, and similar performance to the conventional spot sign on the test set. The 2 features included in the model were: iodine content in the hematoma and iodine content in the brightest spot. These features exemplify the usefulness of quantitative DECT-based features, quantifying contrast medium to aid in accurately capturing the physiology behind hematoma expansion. Future research should focus on gathering more data and implementing more strict protocols to ensure standardization of both input and output.

For the identification and characterization of ischemic lesions on CT acquisitions using deep learning, no generalizable results were obtained. Although the performance of the model on the train set was promising, this appeared to be due to overfitting on the training data since the results could not be generalized to a new, unseen test set. Future research has to focus on the gathering of more data, studying the trade-off between generalization and network complexity in complex deep learning tasks and the standardization of input data.

In the near future, more and more machine learning methods will make their introduction in the hospital, from feature extraction to the classification of medical conditions and the generation of alternative imaging methods. The amount of available data will keep increasing, paving the way for more applications of machine learning in clinical practice.

Bibliography

[1] E. J. Benjamin, M. J. Blaha, S. E. Chiuve, M. Cushman, S. R. Das, R. Deo, S. D. De Ferranti, J. Floyd, M. Fornage, C. Gillespie, C. R. Isasi, M. C. Jimenez, L. C. Jordan, S. E. Judd, D. Lackland, J. H. Lichtman, L. Lisabeth, S. Liu, C. T. Longenecker, R. H. MacKey, K. Matsushita, D. Mozaffarian, M. E. Mussolino, K. Nasir, R. W. Neumar, L. Palaniappan, D. K. Pandey, R. R. Thiagarajan, M. J. Reeves, M. Ritchey, C. J. Rodriguez, G. A. Roth, W. D. Rosamond, C. Sasson, A. Towfghi, C. W. Tsao, M. B. Turner, S. S. Virani, J. H. Voeks, J. Z. Willey, J. T. Wilkins, J.H.Y. Wu, H. M. Alger, S. S. Wong, and P. Muntner. *Heart Disease and Stroke Statistics'2017 Update: A Report from the American Heart Association*, volume 135. 2017. ISBN 0000000000000. doi: 10.1161/CIR.0000000000000485.

[2] J.M. Gebel and J.P. Broderick. Intracerebral hemorrhage. *Neurologic clinics*, 18(2):419–38, 2000. ISSN 0733-8619. doi: 10.1016/S0733-8619(05)70200-0. URL `http://www.ncbi.nlm.nih.gov/pubmed/15087556`.

[3] J.M. Ferro. Update on intracerebral haemorrhage. *Journal of Neurology*, 253(8):985–999, 2006. ISSN 1432-1459. doi: 10.1007/s00415-006-0201-4. URL `https://doi.org/10.1007/s00415-006-0201-4`.

[4] S. Sacco, R. Totaro, D. Toni, C. Marini, D. Cerone, and A. Carolei. Incidence, case-fatalities and 10-year survival of subarachnoid hemorrhage in a population-based registry. *European Neurology*, 62(3):155–160, 2009. ISSN 00143022. doi: 10.1159/000226617.

[5] T. Brott, K. Thalinger, and V. Hertzberg. Hypertension as a risk factor for spontaneous intracerebral hemorrhage. *Stroke*, 17(6):1078–1083, 1986. ISSN 0039-2499. doi: 10.1161/01.STR.17.6.1078. URL `http://stroke.ahajournals.org/cgi/doi/10.1161/01.STR.17.6.1078`.

[6] A.G. Thrift, J.J. McNeil, A. Forbes, and G.A. Donnan. Three important subgroups of hypertensive persons at greater risk of intracerebral hemorrhage. *Hypertension*, 31(6):1223–1229, 1998. ISSN 0194911X. doi: 10.1161/01.HYP.31.6.1223.

[7] C.M. Fisher. Pathological Observations in Hypertensive Cerebral Hemorrhage. *Journal of Neuropathology & Experimental Neurology*, 30(3):536–550, jul 1971. ISSN 0022-3069. URL `http://dx.doi.org/10.1097/00005072-197107000-00015`.

[8] M.I. Aguilar and T.G. Brott. Update in Intracerebral Hemorrhage. *The Neurohospitalist*, 1(3):148–159, 2011. ISSN 19418752. doi: 10.1177/1941875211409050.

[9] S. Kazui, H. Naritomi, H. Yamamoto, T. Sawada, and T. Yamaguchi. Enlargement of spontaneous intracerebral hemorrhage. *Stroke*, 27(10):1783–1787, 1996. ISSN 0039-2499. doi: 10.1161/01.STR.27.10.1783. URL `http://stroke.ahajournals.org/content/27/10/1783`.

[10] J.A. Carter and W. Curry. Intracerebral Hemorrhage: Pathophysiology and Management for Generalists. *Hospital Medicine Clinics*, 6(1):95–111, 2017. ISSN 22115943. doi: 10.1016/j.ehmc.2016.08.002. URL http://dx.doi.org/10.1016/j.ehmc.2016.08.002.

[11] Tmbirkhead. *Brain drawing, edited under Creative Commons license*. Jan 2015. URL https://commons.wikimedia.org/wiki/File:Brain_midsagital_view.png.

[12] T. Brott, J. Broderick, R. Kothari, W. Barsan, T. Tomsick, L. Sauerbeck, J. Spilker, J. Duldner, and J. Khoury. Early hemorrhage growth in patients with intracerebral hemorrhage. *Stroke*, 28(1):1–5, 1997. ISSN 0039-2499.

[13] S.A. Mayer. Recombinant activated factor VII for acute intracerebral hemorrhage. *Stroke*, 38(2 PART 2):763–767, 2007. ISSN 00392499. doi: 10.1161/01.STR.0000254499.46122.22.

[14] S.M. Davis, J. Broderick, M. Hennerici, N.C. Brun, M.N. Diringer, S.A. Mayer, K. Begtrup, and T. Steiner. Hematoma growth is a determinant of mortality and poor outcome after intracerebral hemorrhage. *Neurology*, 66(8):1175–81, 2006. ISSN 0028-3878.

[15] Y. Silva, R. Leira, J. Tejada, J.M. Lainez, J. Castillo, and A. Davalos. Molecular Signatures of Vascular Injury Are Associated With Early Growth of Intracerebral Hemorrhage. *Stroke*, 36(1):86–91, 2005. ISSN 0039-2499. doi: 10.1161/01.STR.0000149615.51204.0b. URL http://stroke.ahajournals.org/cgi/doi/10.1161/01.STR.0000149615.51204.0b.

[16] J.S. Balami and A.M. Buchan. Complications of intracerebral haemorrhage. *The Lancet Neurology*, 11(1):101–118, 2012. ISSN 14744422. doi: 10.1016/S1474-4422(11)70264-2. URL http://dx.doi.org/10.1016/S1474-4422(11)70264-2.

[17] K.E. Wartenberg and S.A. Mayer. Ultra-Early Hemostatic Therapy for Intracerebral Hemorrhage: Future Directions. *Frontiers of Neurology and Neuroscience*, 37:107–129, 2015. ISSN 16622804. doi: 10.1159/000437117.

[18] S.A. Mayer, N.C. Brun, and K. Begtrup. Efficacy and safety of recombinant activated factor VII for acute intracerebral hemorrhage. *England Journal of ...*, pages 2127–2137, 2008. URL http://www.nejm.org/doi/full/10.1056/NEJMoa0707534.

[19] R. Ramnarayan, D. Anto, T.V. Anilkumar, and R. Nayar. Decompressive Hemicraniectomy in Large Putaminal Hematomas: An Indian Experience. *Journal of Stroke and Cerebrovascular Diseases*, 18(1):1–10, 2009. ISSN 10523057. doi: 10.1016/j.jstrokecerebrovasdis.2008.09.001. URL http://dx.doi.org/10.1016/j.jstrokecerebrovasdis.2008.09.001.

[20] C. Fung, M. Murek, W.J. Z'Graggen, A.K. Krähenbühl, O.P. Gautschi, P. Schucht, J. Gralla, K. Schaller, M. Arnold, U. Fischer, H.P. Mattle, A. Raabe, and J. Beck. Decompressive hemicraniectomy in patients with supratentorial intracerebral hemorrhage. *Stroke*, 43(12):3207–3211, 2012. ISSN 00392499. doi: 10.1161/STROKEAHA.112.666537.

[21] S.G. Heuts, S.S. Bruce, B.E. Zacharia, Z.L. Hickman, C.P. Kellner, E.S. Sussman, M.M. McDowell, R.A. Bruce, and E.S. Connolly. Decompressive hemicraniectomy without clot evacuation in dominant-sided intracerebral hemorrhage with ICP crisis. *Neurosurgical Focus*, 34(5):E4, 2013. ISSN 1092-0684. doi: 10.3171/2013.2.FOCUS1326. URL http://thejns.org/doi/10.3171/2013.2.FOCUS1326.

[22] A.D. Mendelow, B.A. Gregson, H.M. Fernandes, G.D. Murray, G.M. Teasdale, D. Terence Hope, A. Karimi, M.D.M. Shaw, and D.H. Barer. Early surgery versus initial conservative treatment in patients with spontaneous supratentorial intracerebral haematomas in the

International Surgical Trial in Intracerebral Haemorrhage (STICH): A randomised trial. *Lancet*, 365(9457):387–397, 2005. ISSN 01406736. doi: 10.1016/S0140-6736(05)17826-X.

[23] L.B. Goldstein and D.L. Simel. Is this patient having a stroke? *Stroke*, 293(19):2391–2402, 2005. doi: 10.1001/jama.293.19.2391. URL `http://stroke.ahajournals.org/cgi/content/abstract/37/9/2378`.

[24] J.B. Fiebach, P.D. Schellinger, A. Gass, T. Kucinski, M. Siebler, A. Villringer, P. Ölkers, J.G. Hirsch, S. Heiland, P. Wilde, O. Jansen, J. Röther, W. Hacke, and K. Sartor. Stroke Magnetic Resonance Imaging Is Accurate in Hyperacute Intracerebral Hemorrhage: A Multicenter Study on the Validity of Stroke Imaging. *Stroke*, 35(2):502–506, 2004. ISSN 00392499. doi: 10.1161/01.STR.0000114203.75678.88.

[25] J.A. Chalela, C.S. Kidwell, L.M. Nentwich, M. Luby, J.A. Butman, A.M. Demchuk, M.D. Hill, N. Patronas, L. Latour, and S. Warach. Magnetic resonance imaging and computed tomography in emergency assessment of patients with suspected acute stroke: a prospective comparison. *The Lancet*, 369(9558):293–298, 2007. ISSN 01406736. doi: 10.1016/S0140-6736(07)60151-2.

[26] T. Brott, J. Broderick, R. Kothari, W. Barsan, T. Tomsick, L. Sauerbeck, J. Spilker, J. Duldner, and J. Khoury. Early hemorrhage growth in patients with intracerebral hemorrhage. *Stroke; a journal of cerebral circulation*, 28:1–5, 1997. ISSN 0039-2499. doi: 10.1161/01.STR.28.1.1.

[27] Y. Fujii, S. Takeuchi, O. Sasaki, T. Minakawa, and R. Tanaka. Multivariate Analysis of Predictors of Hematoma Enlargement in Spontaneous Intracerebral Hemorrhage. *Stroke*, 29(6):1160–1166, 1998. ISSN 0039-2499. doi: 10.1161/01.STR.29.6.1160. URL `http://stroke.ahajournals.org/cgi/doi/10.1161/01.STR.29.6.1160`.

[28] J.P. Broderick, M.N. Diringer, M.D. Hill, N.C. Brun, S.A. Mayer, T. Steiner, B.E. Skolnick, and S.M. Davis. Determinants of intracerebral hemorrhage growth: An exploratory analysis. *Stroke*, 38(3):1072–1075, 2007. ISSN 00392499. doi: 10.1161/01.STR.0000258078.35316.30.

[29] J. N. Goldstein, L.E. Fazen, R. Snider, K. Schwab, S.M. Greenberg, E.E. Smith, M.H. Lev, and J. Rosand. Contrast extravasation on CT angiography predicts hematoma expansion in intracerebral hemorrhage. *Neurology*, 68:889–894, 2007. ISSN 0039-2499.

[30] J.J. Flibotte, N. Hagan, J. O'Donnell, S.M. Greenberg, and J. Rosand. Warfarin, hematoma expansion, and outcome of intracerebral hemorrhage. *Neurology*, 63(6):1059–64, 2004. ISSN 0028-3878.

[31] B. Cucchiara, S. Messe, L. Sansing, S. Kasner, and P. Lyden. Hematoma growth in oral anticoagulant related intracerebral hemorrhage. *Stroke*, 39(11):2993–2996, 2008. ISSN 00392499. doi: 10.1161/STROKEAHA.108.520668.

[32] J.K. Lim, H.S. Hwang, B.M. Cho, H.K. Lee, S.K. Ahn, S.M. Oh, and S.K. Choi. Multivariate analysis of risk factors of hematoma expansion in spontaneous intracerebral hemorrhage. *Surgical Neurology*, 69(1):40–45, 2008. ISSN 00903019. doi: 10.1016/j.surneu.2007.07.025.

[33] R. Takeda, T. Ogura, H. Ooigawa, G. Fushihara, S.I. Yoshikawa, D. Okada, R. Araki, and H. Kurita. A practical prediction model for early hematoma expansion in spontaneous deep ganglionic intracerebral hemorrhage. *Clinical Neurology and Neurosurgery*, 115(7):1028–1031, 2013. ISSN 03038467. doi: 10.1016/j.clineuro.2012.10.016. URL `http://dx.doi.org/10.1016/j.clineuro.2012.10.016`.

[34] D. Rodriguez-Luna, S. Piñeiro, M. Rubiera, M. Ribo, P. Coscojuela, J. Pagola, A. Flores, M. Muchada, B. Ibarra, P. Meler, E. Sanjuan, M. Hernandez-Guillamon, J. Alvarez-Sabin, J. Montaner, and C.A. Molina. Impact of blood pressure changes and course on hematoma growth in acute intracerebral hemorrhage. *European Journal of Neurology*, 20(9):1277–1283, 2013. ISSN 13515101. doi: 10.1111/ene.12180.

[35] A. Morotti, G. Boulouis, J. Romero, H. B. Brouwers, M. J. Jessel, A. Vashkevich, K. Schwab, M.R. Afzal, C. Cassarly, S. Greenberg, R.H. Martin, A. Qureshi, and J. Rosand. Blood pressure reduction and noncontrast CT markers of intracerebral hemorrhage expansion. *Neurology*, 2017. ISSN 2396-9881. doi: http://dx.doi.org/10.1177/2396987317705236. URL `http://ovidsp.ovid.com/ovidweb.cgi?T=JS{&}PAGE=reference{&}D=emexa{&}NEWS=N{&}AN=616967230`.

[36] A. Del Giudice, D. D'Amico, J. Sobesky, and I. Wellwood. Accuracy of the spot sign on computed tomography angiography as a predictor of haematoma enlargement after acute spontaneous intracerebral haemorrhage: A systematic review. *Cerebrovascular Diseases*, 37(4):268–276, 2014. ISSN 14219786. doi: 10.1159/000360754.

[37] A. Morotti, D. Dowlatshahi, G. Boulouis, F. Al-Ajlan, A.M. Demchuk, R.I. Aviv, L. Yu, K. Schwab, J.M. Romero, M.E. Gurol, A. Viswanathan, C.D. Anderson, Y. Chang, S.M. Greenberg, A.I. Qureshi, J. Rosand, J.N. Goldstein, ATACH-II, NETT, and PREDICT Investigators. Predicting Intracerebral Hemorrhage Expansion With Noncontrast Computed Tomography. *Stroke*, page STROKEAHA.117.020138, 2018. ISSN 0039-2499. doi: 10.1161/STROKEAHA.117.020138. URL `http://www.ncbi.nlm.nih.gov/pubmed/29669875{%}0Ahttp://stroke.ahajournals.org/lookup/doi/10.1161/STROKEAHA.117.020138`.

[38] V. Yogendrakumar, A.M. Demchuk, R.I. Aviv, D. Rodriguez-Luna, C.A. Molina, Y.S. Blas, I. Dzialowski, A. Kobayashi, J-M. Boulanger, C. Lum, G. Gubitz, V. Padma, J. Roy, C.S. Kase, R. Bhatia, M.D Hill, and D. Dowlatshahi. Location of intracerebral haemorrhage predicts haematoma expansion. *European Stroke Journal*, 2(3):257–263, 2017. ISSN 2396-9873. doi: 10.1177/2396987317715836. URL `http://journals.sagepub.com/doi/10.1177/2396987317715836`.

[39] C.D. Barras, B.M. Tress, S. Christensen, L. MacGregor, M. Collins, P.M. Desmond, B.E. Skolnick, S.A. Mayer, J.P. Broderick, M.N. Diringer, T. Steiner, and S.M. Davis. Density and shape as CT predictors of intracerebral hemorrhage growth. *Stroke*, 40(4):1325–1331, 2009. ISSN 00392499. doi: 10.1161/STROKEAHA.108.536888.

[40] Q. Li, Q.J. Liu, W.S. Yang, X.C. Wang, L.B. Zhao, X. Xiong, R. Li, D. Cao, D. Zhu, X. Wei, and P. Xie. Island sign: An imaging predictor for early hematoma expansion and poor outcome in patients with intracerebral hemorrhage. *Stroke*, 48(11):3019–3025, 2017. ISSN 15244628. doi: 10.1161/STROKEAHA.117.017985.

[41] Z. Yu, J. Zheng, H. Ali, R. Guo, M. Li, X. Wang, L. Ma, H. Li, and C. You. Significance of satellite sign and spot sign in predicting hematoma expansion in spontaneous intracerebral hemorrhage. *Clinical Neurology and Neurosurgery*, 162(September):67–71, 2017. ISSN 18726968. doi: 10.1016/j.clineuro.2017.09.008.

[42] Z. Yu, J. Zheng, L. Ma, R. Guo, M. Li, X. Wang, S. Lin, H. Li, and C. You. The predictive accuracy of the black hole sign and the spot sign for hematoma expansion in patients with spontaneous intracerebral hemorrhage. *Neurological Sciences*, 38(9):1591–1597, 2017. ISSN 15903478. doi: 10.1007/s10072-017-3006-6.

[43] Q. Li, W-S. Yang, X-C. Wang, D. Cao, D. Zhu, F-J. Lv, Y. Liu, L. Yuan, G. Zhang, X. Xiong, R. Li, Y-X. Hu, and X-Y. Qin. Blend sign predicts poor outcome in patients with intracerebral hemorrhage. *PLoS ONE*, 12(8):e0183082, 2017. ISSN 19326203. doi: http://dx.doi.org/10.1371/journal.pone.0183082. URL `http://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0183082{&}type=printablehttp://ovidsp.ovid.com/ovidweb.cgi?T=JS{&}PAGE=reference{&}D=emexa{&}NEWS=N{&}AN=617933165`.

[44] E. Selariu, E. Zia, M. Brizzi, and K. Abul-Kasim. Swirl sign in intracerebral haemorrhage: definition, prevalence, reliability and prognostic value. *BMC Neurology*, 12(1):1, 2012. ISSN 14712377. doi: 10.1186/1471-2377-12-109.

[45] D. Blacquiere, A.M. Demchuk, M. Al-Hazzaa, A. Deshpande, W. Petrcich, R.I. Aviv, D. Rodriguez-Luna, C.A. Molina, Y. Silva Blas, I. Dzialowski, A. Czlonkowska, J.M. Boulanger, C. Lum, G. Gubitz, V. Padma, J. Roy, C.S. Kase, R. Bhatia, M.D. Hill, and D. Dowlatshahi. Intracerebral Hematoma Morphologic Appearance on Noncontrast Computed Tomography Predicts Significant Hematoma Expansion. *Stroke*, 46(11):3111–3116, 2015. ISSN 15244628. doi: 10.1161/STROKEAHA.115.010566.

[46] G. Boulouis, A. Morotti, H.B. Brouwers, A. Charidimou, M.J. Jessel, E. Auriel, O. Pontes-Neto, A. Ayres, A. Vashkevic, M.E. Gurol, S.M. Greenberg, and J.N. Goldstein. Association between hypodensities detected by Computed Tomography and Hematoma Expansion in Patients With Intracerebral Hemorrhage. *JAMA Neurology*, 73(8):961–968, 2016. ISSN 1527-5418. doi: 10.1097/OGX.0000000000000256.Prenatal.

[47] R. Wada, R.I. Aviv, A.J. Fox, D.J. Sahlas, D.J. Gladstone, G. Tomlinson, and S.P. Symons. CT angiography "spot sign" predicts hematoma expansion in acute intracerebral hemorrhage. *Stroke*, 38(4):1257–1262, 2007. ISSN 00392499. doi: 10.1161/01.STR.0000259633.59404.f3.

[48] Josser E. Delgado Almandoz, Albert J. Yoo, Michael J. Stone, Pamela W. Schaefer, Joshua N. Goldstein, Jonathan Rosand, Alexandra Oleinik, Michael H. Lev, R. Gilberto Gonzalez, and Javier M. Romero. Systematic characterization of the computed tomography angiography spot sign in primary intracerebral hemorrhage identifies patients at highest risk for hematoma expansion: the spot sign score. *Stroke; a journal of cerebral circulation*, 40(9):2994–3000, 2009. ISSN 15244628. doi: 10.1161/STROKEAHA.109.554667.

[49] A.M. Demchuk, D. Dowlatshahi, D. Rodriguez-Luna, C.A. Molina, Y.S. Blas, I. Dzialowski, A. Kobayashi, J.M. Boulanger, C. Lum, G. Gubitz, V. Padma, J. Roy, C.S. Kase, J. Kosior, R. Bhatia, S. Tymchuk, S. Subramaniam, D.J. Gladstone, M.D. Hill, and R.I. Aviv. Prediction of haematoma growth and outcome in patients with intracerebral haemorrhage using the CT-angiography spot sign (PREDICT): A prospective observational study. *The Lancet Neurology*, 11(4):307–314, 2012. ISSN 14744422. doi: 10.1016/S1474-4422(12)70038-8. URL `http://dx.doi.org/10.1016/S1474-4422(12)70038-8`.

[50] A. Tsukabe, Y. Watanabe, H. Tanaka, Y. Kunitomi, M. Nishizawa, A. Arisawa, K. Yoshiya, T. Shimazu, and N. Tomiyama. Prevalence and diagnostic performance of computed tomography angiography spot sign for intracerebral hematoma expansion depend on scan timing. *Neuroradiology*, 56(12):1039–1045, 2014. ISSN 14321920. doi: 10.1007/s00234-014-1430-3.

[51] V.A. Ciura, H.B. Brouwers, R. Pizzolato, C.J. Ortiz, J. Rosand, J.N. Goldstein, S.M. Greenberg, S.R. Pomerantz, R.G. Gonzalez, and J.M. Romero. Spot sign on 90-second delayed computed tomography angiography improves sensitivity for hematoma expansion

and mortality: Prospective study. *Stroke*, 45(11):3293–3297, 2014. ISSN 15244628. doi: 10.1161/STROKEAHA.114.005570.

[52] E. Bandera, M. Botteri, C. Minelli, A. Sutton, K. R. Abrams, and N. Latronico. Cerebral blood flow threshold of ischemic penumbra and infarct core in acute ischemic stroke: A systematic review. *Stroke*, 37(5):1334–1339, 2006. ISSN 00392499. doi: 10.1161/01.STR.0000217418.29609.22.

[53] J-C. Baron and G. Marchal. Ischemic core and penumbra in human stroke. *Stroke*, 30 (5):1150a–1153, 1999. ISSN 0039-2499. doi: 10.1161/01.STR.30.5.1150.a. URL `http://stroke.ahajournals.org/content/30/5/1150a`.

[54] W-D. Heiss. Ischemic penumbra: Evidence from functional imaging in man. *Journal of Cerebral Blood Flow & Metabolism*, 20(9):1276–1293, 2000. doi: 10.1097/00004647-200009000-00002. URL `https://doi.org/10.1097/00004647-200009000-00002`. PMID: 10994849.

[55] J. Astrup, B.K. Siesjö, and L. Symon. Thresholds in cerebral ischemia - the ischemic penumbra. *Stroke*, 12(6):723–725, 1981. ISSN 0039-2499. doi: 10.1161/01.STR.12.6.723. URL `http://stroke.ahajournals.org/content/12/6/723`.

[56] R.G. González, W.A. Copen, P.W. Schaefer, M.H. Lev, S.R. Pomerantz, O. Rapalino, J.W. Chen, G.J. Hunter, J.M. Romero, B.R. Buchbinder, M. Larvie, J.A. Hirsch, and R. Gupta. The Massachusetts General Hospital acute stroke imaging algorithm: An experience and evidence based approach. *Journal of NeuroInterventional Surgery*, 5(SUPPL.1):1–6, 2013. ISSN 17598478. doi: 10.1136/neurintsurg-2013-010715.

[57] S. B. Raymond and P. W. Schaefer. Imaging Brain Collaterals: Quantification, Scoring, and Potential Significance. *Topics in Magnetic Resonance Imaging*, 26(2):67–75, 2017. ISSN 15361004. doi: 10.1097/RMR.0000000000000123.

[58] M. Goyal, B.K. Menon, and C.P. Derdeyn. Perfusion Imaging in Acute Ischemic Stroke: Let Us Improve the Science before Changing Clinical Practice. *Radiology*, 266(1):16–21, 2013. ISSN 0033-8419. doi: 10.1148/radiol.12112134. URL `http://pubs.rsna.org/doi/10.1148/radiol.12112134`.

[59] O.A. Berkhemer, P.S.S. Fransen, D. Beumer, L.A. van den Berg, H.F. Lingsma, A.J. Yoo, W.J. Schonewille, J.A. Vos, P.J. Nederkoorn, M.J.H. Wermer, M.A.A. van Walderveen, J. Staals, J. Hofmeijer, J.A. van Oostayen, .J. Lycklama à Nijeholt, J. Boiten, P.A. Brouwer, B.J. Emmer, S.F. de Bruijn, L.C. van Dijk, L.J. Kappelle, R.H. Lo, E.J. van Dijk, J. de Vries, P.L.M. de Kort, W.J.J. van Rooij, J.S.P. van den Berg, B.A.A.M. van Hasselt, L.A.M. Aerden, R.J. Dallinga, M.C. Visser, J.C.J. Bot, P.C. Vroomen, O. Eshghi, T.H.C.M.L. Schreuder, R.J.J. Heijboer, K. Keizer, A.V. Tielbeek, H.M. den Hertog, D.G. Gerrits, R.M. van den Berg-Vos, G.B. Karas, E.W. Steyerberg, H. Z. Flach, H.A. Marquering, M.E.S. Sprengers, S.F.M. Jenniskens, L.F.M. Beenen, R. van den Berg, P.J. Koudstaal, W.H. van Zwam, Y.B.W.E.M. Roos, A. van der Lugt, R.J. van Oostenbrugge, C.B.L.M. Majoie, and D.W.J. Dippel. A Randomized Trial of Intraarterial Treatment for Acute Ischemic Stroke. *New England Journal of Medicine*, 372(1):11–20, 2015. ISSN 0028-4793. doi: 10.1056/NEJMoa1411587. URL `http://www.nejm.org/doi/10.1056/NEJMoa1411587`.

[60] M. Goyal, A.M. Demchuk, B.K. Menon, M. Eesa, J.L. Rempel, J. Thornton, D. Roy, T.G. Jovin, R.A. Willinsky, B.L. Sapkota, D. Dowlatshahi, D.F. Frei, N.R. Kamal, W.J. Montanera, A.Y. Poppe, K.J. Ryckborst, F.L. Silver, A. Shuaib, D. Tampieri,

D. Williams, O.Y. Bang, B.W. Baxter, P.A. Burns, H. Choe, J. Heo, C.A. Holmstedt, Brian Jankowitz, Michael Kelly, Guillermo Linares, Jennifer L. Mandzia, Jai Shankar, Sung-Il Sohn, Richard H. Swartz, P.A. Barber, S.B. Coutts, E.E. Smith, W.F. Morrish, A. Weill, S. Subramaniam, A.P. Mitha, J.H. Wong, M.W. Lowerison, T.T. Sajobi, and M.D. Hill. Randomized Assessment of Rapid Endovascular Treatment of Ischemic Stroke. *New England Journal of Medicine*, 372(11):1019–1030, 2015. ISSN 0028-4793. doi: 10.1056/NEJMoa1414905. URL `http://www.nejm.org/doi/10.1056/NEJMoa1414905`.

[61] J.P. Broderick, Y.Y. Palesch, A.M. Demchuk, S.D. Yeatts, P. Khatri, M.D. Hill, E.C. Jauch, T.G. Jovin, B. Yan, F.L. Silver, R. von Kummer, C.A. Molina, B.M. Demaerschalk, R. Budzik, W.M. Clark, O.O. Zaidat, T.W. Malisch, M. Goyal, W.J. Schonewille, M. Mazighi, S.T. Engelter, C. Anderson, J. Spilker, J. Carrozzella, K.J. Ryckborst, L.S. Janis, R. H. Martin, L.D. Foster, and T.A. Tomsick. Endovascular Therapy after Intravenous t-PA versus t-PA Alone for Stroke. *New England Journal of Medicine*, 368(10):893–903, 2013. ISSN 0028-4793. doi: 10.1056/NEJMoa1214300. URL `http://www.nejm.org/doi/10.1056/NEJMoa1214300`.

[62] D. Kleindorfer, C.J. Lindsell, L. Brass, W. Koroshetz, and J.P. Broderick. National us estimates of recombinant tissue plasminogen activator use. *Stroke*, 39(3):924–928, 2008. ISSN 0039-2499. doi: 10.1161/STROKEAHA.107.490375. URL `http://stroke.ahajournals.org/content/39/3/924`.

[63] J.A. Zivin. Acute stroke therapy with tissue plasminogen activator (tpa) since it was approved by the u.s. food and drug administration (fda). *Annals of Neurology*, 66(1):6–10, 2009. ISSN 1531-8249. doi: 10.1002/ana.21750. URL `http://dx.doi.org/10.1002/ana.21750`.

[64] A. Ciccone, L. Valvassori, M. Nichelatti, A. Sgoifo, M. Ponzio, R. Sterzi, and E. Boccardi. Endovascular Treatment for Acute Ischemic Stroke. *New England Journal of Medicine*, 368 (10):904–913, 2013. ISSN 0028-4793. doi: 10.1056/NEJMoa1213701. URL `http://www.nejm.org/doi/10.1056/NEJMoa1213701`.

[65] C.S. Kidwell, R. Jahan, J. Gornbein, J.R. Alger, V. Nenov, Z. Ajani, L. Feng, B.C. Meyer, S. Olson, L.H. Schwamm, A.J. Yoo, R.S. Marshall, P.M. Meyers, D.R. Yavagal, M. Wintermark, J. Guzy, S. Starkman, and J.L. Saver. A Trial of Imaging Selection and Endovascular Treatment for Ischemic Stroke. *New England Journal of Medicine*, 368(10):914–923, 2013. ISSN 0028-4793. doi: 10.1056/NEJMoa1212793. URL `http://www.nejm.org/doi/10.1056/NEJMoa1212793`.

[66] J.L. Saver, M. Goyal, A. Bonafe, H.C. Diener, E.I. Levy, V.M. Pereira, G.W. Albers, C. Cognard, D.J. Cohen, W. Hacke, O. Jansen, T.G. Jovin, H.P. Mattle, R.G. Nogueira, A.H. Siddiqui, D.R. Yavagal, B.W. Baxter, T.G. Devlin, D.K. Lopes, V.K. Reddy, R. du Mesnil de Rochemont, O.C. Singer, and R. Jahan. Stent-Retriever Thrombectomy after Intravenous t-PA vs. t-PA Alone in Stroke. *New England Journal of Medicine*, 372(24): 2285–2295, 2015. ISSN 0028-4793. doi: 10.1056/NEJMoa1415061. URL `http://www.nejm.org/doi/10.1056/NEJMoa1415061`.

[67] B.C.V. Campbell, P.J. Mitchell, T.J. Kleinig, H.M. Dewey, L. Churilov, N. Yassi, B. Yan, R.J. Dowling, M.W. Parsons, T.J. Oxley, T.Y. Wu, M. Brooks, M.A. Simpson, F. Miteff, C.R. Levi, M. Krause, T.J. Harrington, K.C. Faulder, B.S. Steinfort, M. Priglinger, T. Ang, R. Scroop, P.A. Barber, B. McGuinness, T. Wijeratne, T.G. Phan, W. Chong, R.V. Chandra, C.F. Bladin, M. Badve, H. Rice, L. de Villiers, H. Ma, P.M. Desmond, G.A. Donnan,

and S.M. Davis. Endovascular Therapy for Ischemic Stroke with Perfusion-Imaging Selection. *New England Journal of Medicine*, 372(11):1009–1018, 2015. ISSN 0028-4793. doi: 10.1056/NEJMoa1414792. URL http://www.nejm.org/doi/10.1056/NEJMoa1414792.

[68] R.G. Nogueira, A.P. Jadhav, D.C. Haussen, A. Bonafe, R.F. Budzik, P. Bhuva, D.R. Yavagal, M. Ribo, C. Cognard, R.A. Hanel, C.A. Sila, A.E. Hassan, M. Millan, E.I. Levy, P. Mitchell, M. Chen, J.D. English, Q.A. Shah, F.L. Silver, V.M. Pereira, B.P. Mehta, B.W. Baxter, M.G. Abraham, P. Cardona, E. Veznedaroglu, F.R. Hellinger, L. Feng, J.F. Kirmani, D.K. Lopes, B.T. Jankowitz, M.R. Frankel, V. Costalat, N.A. Vora, A.J. Yoo, A.M. Malik, A.J. Furlan, M. Rubiera, A. Aghaebrahim, J.M. Olivot, Wo.G. Tekle, R. Shields, T. Graves, R.J. Lewis, W.S. Smith, D.S. Liebeskind, J.L. Saver, and T.G. Jovin. Thrombectomy 6 to 24 Hours after Stroke with a Mismatch between Deficit and Infarct. *New England Journal of Medicine*, page NEJMoa1706442, 2017. ISSN 0028-4793. doi: 10.1056/NEJMoa1706442. URL http://www.nejm.org/doi/10.1056/NEJMoa1706442.

[69] T.M. Leslie-Mazwi, J.A. Hirsch, G.J. Falcone, P.W. Schaefer, M.H. Lev, J.D. Rabinov, N.S. Rost, L. Schwamm, and R.G. González. Endovascular stroke treatment outcomes after patient selection based on magnetic resonance imaging and clinical criteria. *JAMA Neurology*, 73(1):43–49, 2016. ISSN 21686149. doi: 10.1001/jamaneurol.2015.3000.

[70] A.J. Yoo, R. Hu, R. Hakimelahi, M.H. Lev, R.G. Nogueira, J.A. Hirsch, R.G. González, and P.W. Schaefer. CT angiography source images acquired with a fast-acquisition protocol overestimate infarct core on diffusion weighted images in acute ischemic stroke. *Journal of Neuroimaging*, 22(4):329–335, 2012. ISSN 10512284. doi: 10.1111/j.1552-6569.2011.00627.x.

[71] S.F. Zaidi, A. Aghaebrahim, X. Urra, M.A. Jumaa, B. Jankowitz, M. Hammer, R. Nogueira, M. Horowitz, V. Reddy, and T.G. Jovin. Final infarct volume is a stronger predictor of outcome than recanalization in patients with proximal middle cerebral artery occlusion treated with endovascular therapy. *Stroke*, 43(12):3238–3244, 2012. ISSN 0039-2499. doi: 10.1161/STROKEAHA.112.671594. URL http://stroke.ahajournals.org/content/43/12/3238.

[72] A.J. Yoo, L.A. Verduzco, P.W. Schaefer, J.A. Hirsch, J.D. Rabinov, and R.G. González. Mri-based selection for intra-arterial stroke therapy. *Stroke*, 40(6):2046–2054, 2009. ISSN 0039-2499. doi: 10.1161/STROKEAHA.108.541656. URL http://stroke.ahajournals.org/content/40/6/2046.

[73] M.G. Lansberg, M. Straka, S. Kemp, M. Mlynash, L.R. Wechsler, T.G. Jovin, M.J. Wilder, H.L. Lutsep, T.J. Czartoski, R.A. Bernstein, C.W.J. Chang, S. Warach, F. Fazekas, M. Inoue, A. Tipirneni, S.A. Hamilton, G. Zaharchuk, M.P. Marks, R. Bammer, and G.W. Albers. Mri profile and response to endovascular reperfusion after stroke (defuse 2): a prospective cohort study. *The Lancet Neurology*, 11(10):860 – 867, 2012. ISSN 1474-4422. doi: https://doi.org/10.1016/S1474-4422(12)70203-X. URL http://www.sciencedirect.com/science/article/pii/S147444221270203X.

[74] M.G. Lansberg, V.N. Thijs, R. Bammer, S. Kemp, C.A.C. Wijman, M.P. Marks, and G.W. Albers. Risk factors of symptomatic intracerebral hemorrhage after tpa therapy for acute stroke. *Stroke*, 38(8):2275–2278, 2007. ISSN 0039-2499. doi: 10.1161/STROKEAHA.106.480475. URL http://stroke.ahajournals.org/content/38/8/2275.

[75] O.C. Singer, M.C. Humpich, J. Fiehler, G.W. Albers, M.G. Lansberg, A. Kastrup, A. Rovira, D.S. Liebeskind, A. Gass, C. Rosso, L. Derex, J.S. Kim, and T. Neumann-Haefelin. Risk for symptomatic intracerebral hemorrhage after thrombolysis assessed by

diffusion-weighted magnetic resonance imaging. *Annals of Neurology*, 63(1):52–60, 2008. ISSN 1531-8249. doi: 10.1002/ana.21222. URL http://dx.doi.org/10.1002/ana.21222.

[76] M. Wintermark, H.A. Rowley, and M.H. Lev. Acute stroke triage to intravenous thrombolysis and other therapies with advanced ct or mr imaging: Pro ct. *Radiology*, 251(3): 619–626, 2009. doi: 10.1148/radiol.2513081073. URL https://doi.org/10.1148/radiol.2513081073. PMID: 19474369.

[77] R.E. Latchaw, M.J. Alberts, M.H. Lev, J.J. Connors, R.E. Harbaugh, Ra.T. Higashida, R. Hobson, C.S. Kidwell, W.J. Koroshetz, V. Mathews, P. Villablanca, S. Warach, and B. Walters. Recommendations for imaging of acute ischemic stroke. *Stroke*, 40(11):3646–3678, 2009. ISSN 0039-2499. doi: 10.1161/STROKEAHA.108.192616. URL http://stroke.ahajournals.org/content/40/11/3646.

[78] J.V. Guadagno, E.A. Warburton, P.S. Jones, T.D. Fryer, D.J. Day, J.H. Gillard, T.A. Carpenter, F.I. Aigbirhio, C.J. Price, and J. Baron. The diffusion-weighted lesion in acute stroke: Heterogeneous patterns of flow/metabolism uncoupling as assessed by quantitative positron emission tomography. *Cerebrovascular Disease*, pages 239–246, 2005.

[79] K.O. Lovblad, H.J. Laubach, A.E. Baird, F. Curtin, G. Schlaug, R.R. Edelman, and S. Warach. Clinical experience with diffusion-weighted mr in patients with acute stroke. *American Journal of Neuroradiology*, 19(6):1061–1066, 1998. ISSN 0195-6108. URL http://www.ajnr.org/content/19/6/1061.

[80] K.J. van Everdingen, J. van der Grond, L.J. Kappelle, L.M.P. Ramos, and W.P.T.M. Mali. Diffusion-weighted magnetic resonance imaging in acute stroke. *Stroke*, 29(9):1783–1790, 1998. ISSN 0039-2499. doi: 10.1161/01.STR.29.9.1783. URL http://stroke.ahajournals.org/content/29/9/1783.

[81] M.E. Mullins, P.W. Schaefer, A.G. Sorensen, E.F. Halpern, H. Ay, J. He, W.J. Koroshetz, and R.G. Gonzalez. Ct and conventional and diffusion-weighted mr imaging in acute stroke: Study in 691 patients at presentation to the emergency department. *Radiology*, 224(2):353–360, 2002. doi: 10.1148/radiol.2242010873. URL https://doi.org/10.1148/radiol.2242010873. PMID: 12147827.

[82] W.J. Powers, C.P. Derdeyn, J. Biller, C.S. Coffey, B.L. Hoh, E.C. Jauch, K.C. Johnston, S. C. Johnston, A.A. Khalessi, C.S. Kidwell, J.F. Meschia, B. Ovbiagele, and D.R. Yavagal. 2015 aha/asa focused update of the 2013 guidelines for the early management of patients with acute ischemic stroke regarding endovascular treatment. *Stroke*, 2015. ISSN 0039-2499. doi: 10.1161/STR.0000000000000074. URL http://stroke.ahajournals.org/content/early/2015/06/26/STR.0000000000000074.

[83] K.W. Muir, A. Buchan, R. von Kummer, J. Rother, and J-C. Baron. Imaging of acute stroke. *The Lancet Neurology*, 5(9):755 – 768, 2006. ISSN 1474-4422. doi: https://doi.org/10.1016/S1474-4422(06)70545-2. URL http://www.sciencedirect.com/science/article/pii/S1474442206705452.

[84] W. Hacke, M. Kaste, and C. Fieschi. Intravenous thrombolysis with recombinant tissue plasminogen activator for acute hemispheric stroke: The european cooperative acute stroke study (ecass). *JAMA*, 274(13):1017–1025, 1995. doi: 10.1001/jama.1995.03530130023023. URL +http://dx.doi.org/10.1001/jama.1995.03530130023023.

[85] F. Miteff, C.R. Levi, .A. Bateman, N. Spratt, P. McElduff, and M.W. Parsons. The independent predictive utility of computed tomography angiographic collateral status in acute ischaemic stroke. *Brain*, 132(8):2231–2238, 2009. doi: 10.1093/brain/awp155. URL +http://dx.doi.org/10.1093/brain/awp155.

[86] M.B. Maas, M.H. Lev, H. Ay, A.B. Singhal, D.M. Greer, W.S. Smith, G.J. Harris, E. Halpern, A. Kemmling, W.J. Koroshetz, and K.L. Furie. Collateral vessels on ct angiography predict outcome in acute ischemic stroke. *Stroke*, 40(9):3001–3005, 2009. ISSN 0039-2499. doi: 10.1161/STROKEAHA.109.552513. URL http://stroke.ahajournals.org/content/40/9/3001.

[87] L.C.S. Souza, A.J. Yoo, Z.A. Chaudhry, S. Payabvash, A. Kemmling, P.W. Schaefer, J.A. Hirsch, K.L. Furie, R.G. González, R.G. Nogueira, and M.H. Lev. Malignant cta collateral profile is highly specific for large admission dwi infarct core and poor outcome in acute stroke. *American Journal of Neuroradiology*, 33(7):1331–1336, 2012. ISSN 0195-6108. doi: 10.3174/ajnr.A2985. URL http://www.ajnr.org/content/33/7/1331.

[88] E.J. Smit, E-J. Vonken, I.C. van der Schaaf, A.M. Mendrik, J.W. Dankbaar, A.D. Horsch, T. van Seeters, B. van Ginneken, and M. Prokop. Timing-invariant reconstruction for deriving high-quality ct angiographic data from cerebral ct perfusion data. *Radiology*, 263(1):216–225, 2012. doi: 10.1148/radiol.11111068. URL https://doi.org/10.1148/radiol.11111068. PMID: 22332063.

[89] B.K. Menon, B. O'Brien, A. Bivard, N.J. Spratt, A.M. Demchuk, F. Miteff, X. Lu, C. Levi, and M.W. Parsons. Assessment of leptomeningeal collaterals using dynamic ct angiography in patients with acute ischemic stroke. *Journal of Cerebral Blood Flow & Metabolism*, 33 (3):365–371, 2013. doi: 10.1038/jcbfm.2012.171. URL https://doi.org/10.1038/jcbfm.2012.171. PMID: 23149554.

[90] J.C. Tan, W.P. Dillon, S. Liu, F. Adler, W.S. Smith, and M. Wintermark. Systematic comparison of perfusion-ct and ct-angiography in acute stroke patients. *Annals of Neurology*, 61(6):533–543, 2007. ISSN 1531-8249. doi: 10.1002/ana.21130. URL http://dx.doi.org/10.1002/ana.21130.

[91] P. Schramm, P.D. Schellinger, E. Klotz, K. Kallenberg, J.B. Fiebach, S. Külkens, S. Heiland, M. Knauth, and K. Sartor. Comparison of perfusion computed tomography and computed tomography angiography source images with perfusion-weighted imaging and diffusion-weighted imaging in patients with acute stoke of less than 6 hours' duration. *Stroke*, 35(7):1652–1657, 2004. ISSN 00392499. doi: 10.1161/01.STR.0000131271.54098.22.

[92] V.N. Vapnik. *The Nature of Statistical Learning Theory*. Springer-Verlag, Berlin, Heidelberg, 1995. ISBN 0-387-94559-8.

[93] V.N. Vapnik. An overview of statistical learning theory. *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council*, 10(5):988–99, 1999. ISSN 1045-9227. doi: 10.1109/72.788640. URL http://www.ncbi.nlm.nih.gov/pubmed/18252602.

[94] B. Schölkopf and A.J. Smola. *Learning with kernels*. MIT Press, 2001. ISBN 9780262194754.

[95] Y. Lecun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015. ISSN 14764687. doi: 10.1038/nature14539.

[96] O. Ronneberger, P. Fischer, and T. Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. pages 1–8, 2015. ISSN 16113349. doi: 10.1007/978-3-319-24574-4_28. URL http://arxiv.org/abs/1505.04597.

[97] Y.L. Cun, L. Bottou, Y. Bengio, and P. Haffiner. Gradient based learning applied to document recognition. *Proceedings of IEEE*, 86(11):86(11):2278–2324, 1998.

[98] S.A. Mayer, S.M. Davis, B.E. Skolnick, N.C. Brun, K. Begtrup, J.P. Broderick, M.N. Diringer, and T. Steiner. Can a subset of intracerebral hemorrhage patients benefit from hemostatic therapy with recombinant activated factor VII? *Stroke*, 40(3):833–840, 2009. ISSN 00392499. doi: 10.1161/STROKEAHA.108.524470.

[99] A.I. Qureshi, Y.Y. Palesch, W.G. Barsan, D.F. Hanley, C.Y. Hsu, R.L. Martin, C.S. Moy, R. Silbergleit, T. Steiner, J.I. Suarez, K. Toyoda, Y. Wang, H. Yamamoto, and B-W. Yoon. Intensive Blood-Pressure Lowering in Patients with Acute Cerebral Hemorrhage. *New England Journal of Medicine*, 375(11):1033–1043, 2016. ISSN 0028-4793. doi: 10.1056/NEJMoa1603460. URL `http://www.nejm.org/doi/10.1056/NEJMoa1603460`.

[100] C.S. Anderson, E. Heeley, Y. Huang, J. Wang, C. Stapf, C. Delcourt, R. Lindley, T. Robinson, P. Lavados, B. Neal, J. Hata, H. Arima, M. Parsons, Yuechun Li, Jinchao Wang, Stephane Heritier, Qiang Li, Mark Woodward, R. John Simes, Stephen M. Davis, and John Chalmers. Rapid Blood-Pressure Lowering in Patients with Acute Intracerebral Hemorrhage. *New England Journal of Medicine*, 368(25):2355–2365, 2013. ISSN 0028-4793. doi: 10.1056/NEJMoa1214609. URL `http://www.nejm.org/doi/10.1056/NEJMoa1214609`.

[101] R. Gupta, C.M. Phan, C. Leidecker, T.J. Brady, J.A Hirsch, R.G. Nogueira, and A.J. Yoo. Evaluation of dual-energy CT for differentiating intracerebral hemorrhage from iodinated contrast material staining. *Radiology*, 257(1):205–211, 2010. ISSN 0033-8419. doi: 10.1148/radiol.10091806.

[102] C.M. Phan, A.J. Yoo, J.A. Hirsch, R.G. Nogueira, and R. Gupta. Differentiation of hemorrhage from iodinated contrast in different intracranial compartments using dual-energy head CT. *American Journal of Neuroradiology*, 33(6):1088–1094, 2012. ISSN 01956108. doi: 10.3174/ajnr.A2909.

[103] J. Yuan, E. Bae, and X.C. Tai. A study on continuous max-flow and min-cut approaches. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 2217–2224. ISBN 1063-6919. doi: 10.1109/CVPR.2010.5539903.

[104] J. Yuan, E. Bae, X. Tai, and Y. Boykov. *A Continuous Max-Flow Approach to Potts Model*, pages 379–392. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010. ISBN 978-3-642-15567-3.

[105] J. Yuan, E. Bae, X. Tai, and Y. Boykov. A spatially continuous max-flow and min-cut framework for binary labeling problems. *Numerische Mathematik*, 126(3):559–587, 2014. ISSN 0029-599X.

[106] C.O. Tan, J.A. Taylor, A.S.H. Ler, and M.A. Cohen. Detection of multifiber neuronal firings: A mixture separation model applied to sympathetic recordings. *IEEE Transactions on Biomedical Engineering*, 56(1):147–158, Jan 2009. ISSN 0018-9294. doi: 10.1109/TBME.2008.2002138.

[107] R.M. Haralick, K. Shanmugam, and I. Dinstein. Textural Features for Image Classification. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-3(6):610–621, 1973.

[108] D.J. Beal. Information Criteria Methods in SAS for Multiple Linear Regression Models. *SESUG Proceedings*, (c):1–10, 2013. URL `http://analytics.ncsu.edu/sesug/2007/SA05.pdf`.

[109] J.M. Romero, H.B. Brouwers, J. Lu, J.E. Delgado Almandoz, H. Kelly, J. Heit, J. Goldstein, J. Rosand, and R.G. Gonzalez. Prospective validation of the computed tomographic angiography spot sign score for Intracerebral hemorrhage. *Stroke*, 44(11):3097–3102, 2013. ISSN 00392499. doi: 10.1161/STROKEAHA.113.002752.

[110] A. Morotti, J.M. Romero, M.J. Jessel, H.B. Brouwers, R. Gupta, K. Schwab, A. Vashkevich, A. Ayres, C.D. Anderson, M.E. Gurol, A. Viswanathan, S.M. Greenberg, J. Rosand, and J.N. Goldstein. Effect of CTA tube current on spot sign detection and accuracy for prediction of intracerebral hemorrhage expansion. *American Journal of Neuroradiology*, 37 (10):1781–1786, 2016. ISSN 1936959X. doi: 10.3174/ajnr.A4810.

[111] K. Orito, M. Hirohata, Y. Nakamura, N. Takeshige, T. Aoki, G. Hattori, K. Sakata, T. Abe, Y. Uchiyama, T. Sakamoto, and M. Morioka. Leakage sign for primary intracerebral hemorrhage a novel predictor of hematoma growth. *Stroke*, 47(4):958–963, 2016. ISSN 15244628. doi: 10.1161/STROKEAHA.115.011578.

[112] N. Liu, H. Chen, B. Wu, Y. Li, M. Wintermark, A. Jackson, J. Hu, Y. Zhang, Z. Su, G. Zhu, and W. Zhang. CT Permeability Imaging Predicts Clinical Outcomes in Acute Ischemic Stroke Patients Treated with Intra-arterial Thrombolytic Therapy. *Molecular Neurobiology*, 54(4):2539–2546, 2017. ISSN 15591182. doi: 10.1007/s12035-016-9838-x.

[113] Q. Li, W-S. Yang, X-C. Wang, D. Cao, D. Zhu, F-J. Lv, Y. Liu, L. Yuan, G. Zhang, X. Xiong, R. Li, Y-X. Hu, and X-Y. Qin. Blend sign predicts poor outcome in patients with intracerebral hemorrhage. *PLoS ONE*, 12(8):e0183082, 2017. ISSN 19326203. doi: http://dx.doi.org/10.1371/journal.pone.0183082. URL `http://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0183082{&}type=printable{%}0Ahttp://ovidsp.ovid.com/ovidweb.cgi?T=JS{&}PAGE=reference{&}D=emexa{&}NEWS=N{&}AN=617933165`.

[114] Q. Shen, Y. Shan, Z. Hu, W. Chen, B. Yang, J. Han, Y. Huang, and W. Xu. Quantitative parameters of CT texture analysis as potential markers for early prediction of spontaneous intracranial hemorrhage enlargement. *European radiology*, pages 1–8, 2018.

[115] A. Ederies, A. Demchuk, T. Chia, D.J. Gladstone, D. Dowlatshahi, G. Bendavit, K. Wong, S.P. Symons, and R.I. Aviv. Postcontrast CT extravasation is associated with hematoma expansion in cta spot negative patients. *Stroke*, 40(5):1672–1676, 2009. ISSN 00392499. doi: 10.1161/STROKEAHA.108.541201.

[116] G. Litjens, T. Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, J.A.W.M. van der Laak, B. van Ginneken, and C.I. Sánchez. A Survey on Deep Learning in Medical Image Analysis. (1995), 2017. ISSN 1361-8423. doi: 10.1016/j.media.2017.07.005. URL `http://arxiv.org/abs/1702.05747http://dx.doi.org/10.1016/j.media.2017.07.005`.

[117] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, and J. Liang. Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning? *IEEE Transactions on Medical Imaging*, 35(5):1299–1312, 2016. ISSN 1558254X. doi: 10.1109/TMI.2016.2535302.

[118] S. Ioffe and C. Szegedy. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. 2015. ISSN 0717-6163. doi: 10.1007/s13398-014-0173-7.2. URL `http://arxiv.org/abs/1502.03167`.

[119] A.E. Orhan and X. Pitkow. Skip Connections Eliminate Singularities. 2017. URL http://arxiv.org/abs/1701.09175.

[120] M. Drozdzal, E. Vorontsov, G. Chartrand, S. Kadoury, and C. Pal. The importance of skip connections in biomedical image segmentation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10008 LNCS:179–187, 2016. ISSN 16113349. doi: 10.1007/978-3-319-46976-8_19.

[121] C.H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. Jorge Cardoso. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10553 LNCS:240–248, 2017. ISSN 16113349. doi: 10.1007/978-3-319-67558-9_28.

[122] C. Zhang, S. Bengio, M. Hardt, B. Recht, and O. Vinyals. Understanding deep learning requires rethinking generalization. 2016.

[123] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.*, 15(1): 1929–1958, January 2014. ISSN 1532-4435. URL http://dl.acm.org/citation.cfm?id=2627435.2670313.