



# Improved Flexibility and Interpretability of Large Vessel Stroke Prognostication Using Image Synthesis and Multi-task Learning

Minyan Zeng<sup>1,2(✉)</sup>, Yutong Xie<sup>1</sup>, Minh-Son To<sup>1,3</sup>, Lauren Oakden-Rayner<sup>1,4</sup>, Luke Whitbread<sup>1,5</sup>, Stephen Bacchi<sup>3,4</sup>, Alix Bird<sup>1,2</sup>, Luke Smith<sup>1,2</sup>, Rebecca Scroop<sup>4</sup>, Timothy Kleinig<sup>4</sup>, Jim Jannes<sup>4</sup>, Lyle J Palmer<sup>1,2</sup>, and Mark Jenkinson<sup>1,5,6</sup>

<sup>1</sup> Australian Institute for Machine Learning, University of Adelaide, Adelaide, Australia

[minyan.zeng@adelaide.edu.au](mailto:minyan.zeng@adelaide.edu.au)

<sup>2</sup> School of Public Health, University of Adelaide, Adelaide, Australia

<sup>3</sup> Flinders Health and Medical Research Institute, Flinders University, Bedford Park, Australia

<sup>4</sup> Royal Adelaide Hospital, Adelaide, Australia

<sup>5</sup> School of Computer and Mathematical Sciences, University of Adelaide, Adelaide, Australia

<sup>6</sup> South Australian Health and Medical Research Institute, Adelaide, Australia

**Abstract.** While acute ischemic stroke due to large vessel occlusion (LVO) may be life-threatening or permanently disabling, timely intervention with endovascular thrombectomy (EVT) can prove life-saving for affected patients. Appropriate patient selection based on prognostic prediction is vital for this costly and invasive procedure, as not all patients will benefit from EVT. Accurate prognostic prediction for LVO presents a significant challenge. Computed Tomography Perfusion (CTP) maps can provide additional information for clinicians to make decisions. However, CTP maps are not always available due to variations in available equipment, funding, expertise and image quality. To address these gaps, we test (i) the utility of acquired CTP maps in a deep learning prediction model, (ii) the ability to improve flexibility of this model through image synthesis, and (iii) the added benefits of including multi-task learning with a simple clinical task to focus the synthesis on key clinical features. Our results demonstrate that network architectures utilising a full set of images can still be flexibly deployed if CTP maps are unavailable as their benefits can be effectively synthesized from more widely available images (NCCT and CTA). Additionally, such synthesized images may help with interpretability and building a clinically trusted model.

---

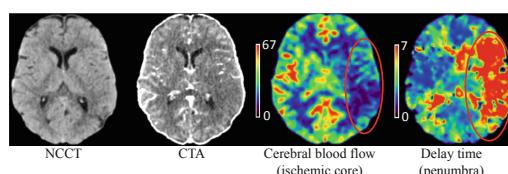
**Supplementary Information** The online version contains supplementary material available at [https://doi.org/10.1007/978-3-031-43904-9\\_67](https://doi.org/10.1007/978-3-031-43904-9_67).

## 1 Introduction

Ischemic stroke caused by large vessel occlusion (LVO) is one of the leading causes of death and disability worldwide [1]. These occlusions restrict cerebral blood supply, resulting in rapid and irreversible damage to brain tissues. As such, prompt reperfusion treatment is critical to restore blood flow and salvage brain tissues from permanent damage. Endovascular thrombectomy (EVT) is the standard treatment recommended for LVO patients [2]. However, this treatment is costly and invasive and does not improve prognoses for all LVO patients [3]. Therefore, it is vital to identify those most likely to benefit from EVT.

There have been a number of models proposed for the prognostication of LVO in recent years [4–6]. The majority are traditional statistical models using non-imaging data, with the identification of clinical predictors for prognostication having shown limited capability [7]. A few models incorporate raw image data using deep-learning techniques, while most of these [4,5] utilize a single image modality, such as Non-Contrast Computed Tomography (NCCT).

CT Perfusion (CTP) maps are increasingly utilized for the prediction of LVO outcomes because they offer quantitative information on blood flow and volume in the brain, as well as the arrival time of blood bolus to brain tissues [8]. This information can show the brain perfusion status and identify the brain tissues that are irreversibly damaged (ischemic core), and those that are potentially salvageable (penumbra) [8]. In contrast, NCCT images can highlight the ischemic core as these tissues normally appear hypodense and CT angiography (CTA) can illustrate the collateral supply via highlighting vessel structure using a low dose contrast injection [9]. These two modalities are routinely collected at patients' admission because of their rapid acquisition and high tolerance [2]. Although the information estimated by NCCT and CTA can reflect the status of blood perfusion to some extent [9], these images are not as clinically useful as CTP maps since they do not directly show the most prominent neuropathological changes of ischemic stroke—ischemic core and salvageable penumbra. An example of ischemic changes shown by NCCT, CTA and CTP maps is presented in Fig. 1, where it can be seen that CTP maps provide clearer visual information, and hence are likely to lead to more accurate and clinically trusted prediction models. As such, it is important to investigate the benefits of incorporating CTP maps into deep learning prediction models.



**Fig. 1.** Examples of NCCT, CTA, and two CTP maps

Despite the advantages of CTP maps, several factors have limited their utilization: (i) it is time-consuming to acquire, process and interpret the images [2]; (ii) poor quality maps are more likely due to motion artefacts and inadequate contrast agent, potentially making them uninterpretable [10]; and (iii) they require substantial investment to purchase advanced scanners, hire professional staff to run and maintain the scanner, which limits usage in hospitals with limited funding, such as in many rural and remote areas [11]. Therefore, it will be helpful if the flexibility of models incorporating CTP maps can be improved through image synthesis from commonly available image modalities (*e.g.*, NCCT and CTA). In this way, patients without access to a CTP scanner, or with uninterpretable CTP maps, can still benefit from the prognostic prediction that uses clinically-relevant features of CTP maps to inform treatment selection.

In recent years, techniques of image synthesis have shown promising potential in medicine [12–14]. However, most of these models have focused on image-to-image conversion tasks. In this paper, we propose a two-stage model that incorporates clinical knowledge for image synthesis and multimodal prognostic prediction. Specifically, in the image synthesis stage, the model was assigned to optimize a joint task, including a generative task, a discriminative task and a clinically-focused task. In the multimodal prognostic prediction stage, the model utilized imaging and non-imaging clinical information for predicting the dichotomised modified Rankin Scale (mRS) score 3 months post-stroke — the main outcome in stroke prognosis.

## 2 Method

### 2.1 Dataset and Pre-processing

Data utilized in this research was collected from the Royal Adelaide Hospital, which provides the sole EVT service to all stroke patients in South Australia and Northern Territory. There were 460 LVO patients included in the study, admitted between 01 Dec 2016 and 01 Dec 2021, and treated with EVT with full image modalities (NCCT, CTA and CTP maps). Of these, 256 achieved functional independency (mRS  $\leq 2$ ) 3 months post-stroke. The non-imaging data (*i.e.*, age, stroke severity, blood glucose, pre-admission functional status, use of intravenous thrombolysis, onset-to-groin puncture time, stroke hemisphere, and 3-month mRS score) was collected by experienced neurologists and nurses adhering to the standard admission procedure. The study was approved by The Central Adelaide Local Health Network Human Research Ethics Committee.

The NCCT and CTA images were skull-stripped with the attenuation clipped between 0 and 100 Hounsfield Units (HU) for the NCCT images and 0 and 750 HU for the CTA images. Multimodal CT imaging data, including NCCT, CTA and CTP maps, were acquired using Canon Aquilion ONE scanners. The NCCT and CTA acquisitions have isotropic voxel sizes ranging from 0.4–0.6 mm and 0.4–0.7 mm, respectively. The acquisition voxel size of the CTP maps is  $0.4 \times 0.4 \times 4.9 \text{ mm}^3$ . Four CTP maps, including cerebral blood volume (CBV), cerebral blood flow (CBF), mean transit time (MTT), and relative arrival time

of contrast (Delay), were selected for their clinical utility, based on consultations with two senior neurologists. To rule out the impact of different brain sizes, affine registration to a CT template was performed for each modality [15, 16]. The CTP maps were linearly scaled versions of the “true” (quantitative) maps using the within-scan relative values for prognostic prediction, rather than absolute values that may be influenced by a range of nuisance factors (e.g. head size, blood pressure) [17]. The image intensities for each modality were normalised to the interval of [0,1] by rescaling the min-max range for prognostication. Images were resampled to a 1mm isotropic resolution.

## 2.2 Models

**Problem Statement.** We take  $\{\mathbf{X}_{\text{NCCT}}^i, \mathbf{X}_{\text{CTA}}^i, \mathbf{X}_{\text{CTP}}^i, \mathbf{X}_{\text{Cli\_var}}^i, y_{\text{mRS}}^i\}_{i=1}^N$  to be a set of data for  $N$  patients, where  $\mathbf{X}_{\text{NCCT}}^i$ ,  $\mathbf{X}_{\text{CTA}}^i$ , and  $\mathbf{X}_{\text{Cli\_var}}^i$  are the NCCT, CTA and clinical non-imaging data (*i.e.*, age, stroke severity, blood glucose, pre-admission functional status, use of intravenous thrombolysis, and onset-to-groin puncture time) for the  $i^{\text{th}}$  patient. Four CTP maps for the  $i^{\text{th}}$  patient are denoted by the set  $\mathbf{X}_{\text{CTP}}^i$  which is defined as  $\{\mathbf{X}_{\text{CBF}}^i, \mathbf{X}_{\text{CBV}}^i, \mathbf{X}_{\text{MTT}}^i, \mathbf{X}_{\text{Delay}}^i\}$ . The dichotomized prognostic outcome for the  $i^{\text{th}}$  patient is denoted by  $y_{\text{mRS}}^i$  defined as {0 if  $mRS \leq 2$ , or 1 if  $mRS > 2$ }.

We aim to: (i) evaluate the performance of CTP maps in predicting the dichotomized mRS score; and (ii) synthesize the CTP maps using two commonly used image modalities (NCCT and CTA) at admission for prognostic prediction. For the first aim, the model can be written as:

$$\hat{y}_{\text{acq}}^i = F_{\text{acq}}(\mathbf{X}_{\text{CBF}}^i, \mathbf{X}_{\text{CBV}}^i, \mathbf{X}_{\text{MTT}}^i, \mathbf{X}_{\text{Delay}}^i, \mathbf{X}_{\text{Cli\_var}}^i) \quad (1)$$

where  $\hat{y}_{\text{acq}}^i$  is the predicted outcome and  $F_{\text{acq}}$  is the mapping function from the acquired CTP maps and clinical information to the dichotomized mRS score.

For the second aim, there are two tasks, including (i) learning a mapping function  $G$  for CTP map generation, and (ii) learning a function  $F_{\text{syn}}$  to map synthetic CTP maps and clinical information to the dichotomized mRS score. That is:

$$(\hat{\mathbf{X}}_{\text{CBF}}^i, \hat{\mathbf{X}}_{\text{CBV}}^i, \hat{\mathbf{X}}_{\text{MTT}}^i, \hat{\mathbf{X}}_{\text{Delay}}^i) = G(\mathbf{X}_{\text{NCCT}}^i, \mathbf{X}_{\text{CTA}}^i), \quad (2)$$

$$\hat{y}_{\text{syn}}^i = F_{\text{syn}}(\hat{\mathbf{X}}_{\text{CBF}}^i, \hat{\mathbf{X}}_{\text{CBV}}^i, \hat{\mathbf{X}}_{\text{MTT}}^i, \hat{\mathbf{X}}_{\text{Delay}}^i, \mathbf{X}_{\text{Cli\_var}}^i) \quad (3)$$

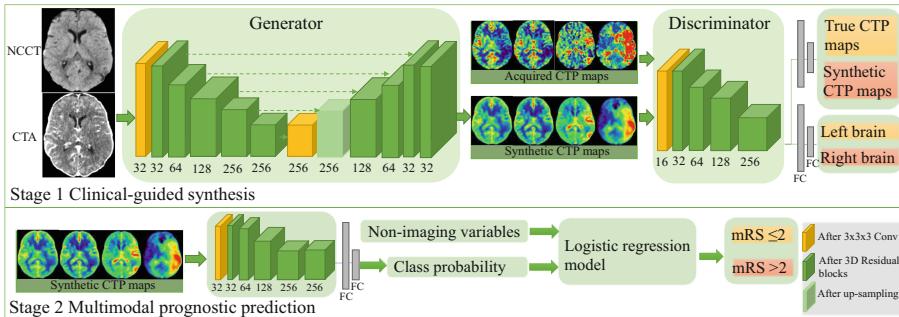
where  $\hat{\mathbf{X}}_{\text{CBF}}^i, \hat{\mathbf{X}}_{\text{CBV}}^i, \hat{\mathbf{X}}_{\text{MTT}}^i, \hat{\mathbf{X}}_{\text{Delay}}^i$  are the predicted CBF, CBV, MTT, and Delay maps and  $\hat{y}_{\text{syn}}^i$  is the predicted outcome from synthetic CTP maps for the  $i^{\text{th}}$  patient. To fulfil the second aim, we propose a two-stage deep learning framework, including a clinical-guided synthesis and a multimodal prognostic prediction. The network architecture and loss function are detailed below.

**Stage 1: Clinical-Guided Synthesis.** The method for synthesizing CTP maps utilizes a 3D generative adversarial network (GAN) model, which is illustrated in Fig. 2 - stage 1. The encoder-decoder is a U-net architecture [18]. The encoder

contains one convolutional layer (32 filters) and five 3D-residual blocks (32, 64, 128, 256, 256 filters) with  $3 \times 3 \times 3$  kernels. Each convolutional layer in the residual block is followed by instance normalization and LeakyReLU activation. The decoder contains five 3D-residual blocks (128, 64, 32, 32, 4 filters) with  $3 \times 3 \times 3$  kernels. After each residual block, features were upsampled and combined with encoder outputs, as usual. NCCT and CTA inputs were concatenated at the channel level. The discriminator contains one convolutional layer (16 filters) and four 3D-residual blocks (32, 64, 128, 256 filters). Real CTP maps and synthetic CTP maps (*e.g.*,  $\mathbf{X}_{\text{CBF}}^i$  and  $\hat{\mathbf{X}}_{\text{CBF}}^i$ ) were input to the discriminator, where two classification heads were designed for: a discriminative task with four fully connected (FC) layers (filters from 128 to 2) to distinguish real from synthetic maps; and a clinical task with a FC layer (filters from 128 to 2) to identify the cerebral hemisphere of the occlusion. The loss function is:

$$\begin{aligned} \text{Loss}_{\text{Stage1}} = & \text{Loss}_{\text{mse}}(G(\mathbf{X}_{\text{NCCT}}^i, \mathbf{X}_{\text{CTA}}^i), \mathbf{X}_{\text{CTP}}^i) \\ & - [\text{Loss}_{\text{bce}}(D(G(\mathbf{X}_{\text{NCCT}}^i, \mathbf{X}_{\text{CTA}}^i)), 0) + \text{Loss}_{\text{bce}}(D(\mathbf{X}_{\text{CTP}}^i), 1)] \\ & - [\text{Loss}_{\text{bce}}(D(G(\mathbf{X}_{\text{NCCT}}^i, \mathbf{X}_{\text{CTA}}^i)), y_{\text{hemi}}^i) + \text{Loss}_{\text{bce}}(D(\mathbf{X}_{\text{CTP}}^i), y_{\text{hemi}}^i)] \end{aligned} \quad (4)$$

where  $\text{Loss}_{\text{mse}}$  and  $\text{Loss}_{\text{bce}}$  calculate the mean square error and the binary cross entropy, respectively.  $D$  is a mapping function for discriminative and clinical tasks.  $y_{\text{hemi}}^i$  is the label of the clinical task {0 : occlusion in the left hemispere; 1 : occlusion in the right hemispere}. We used the total loss for the set of synthetic CTP maps for backpropagation.



**Fig. 2.** Overview of the proposed method for image synthesis and prognostic prediction

**Stage 2: Multimodal Prognostic Prediction.** In the second stage, we used synthetic CTP images  $\hat{\mathbf{X}}_{\text{CBF}}^i, \hat{\mathbf{X}}_{\text{CBV}}^i, \hat{\mathbf{X}}_{\text{MTT}}^i, \hat{\mathbf{X}}_{\text{Delay}}^i$  from stage 1 incorporating non-imaging variables to predict the dichotomized mRS score. The architecture is illustrated in Fig. 2 - stage 2. Specifically, the first step is to train a model with synthetic image inputs. The model contains one convolutional layer (32 filters), five 3D-residual blocks (32, 64, 128, 256, 256 filters) with  $3 \times 3 \times 3$  kernels and a FC layer (filters from 128 to 2). The second step is to train a

logistic regression model (single node with sigmoid activation) with the inputs including the predicted class probability derived from the first step and non-imaging variables. The loss functions of stage 2 can be formulated as:

$$\text{Step 1 : } Loss_{\text{img}} = Loss_{\text{bce}} \left( F_{\text{img}} \left( \hat{\mathbf{X}}_{\text{CTP}}^i \right), y_{\text{mRS}}^i \right) \quad (5)$$

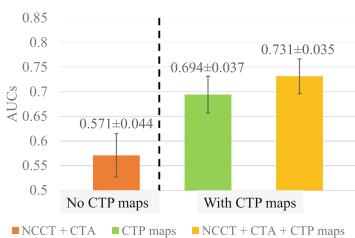
$$\text{Step 2 : } Loss_{\text{logistic}} = Loss_{\text{bce}} \left( F_{\text{logistic}} \left( F_{\text{img}} \left( \hat{\mathbf{X}}_{\text{CTP}}^i \right), \mathbf{X}_{\text{Cli\_var}}^i \right), y_{\text{mRS}}^i \right) \quad (6)$$

where  $F_{\text{img}}$  is a mapping function of  $\hat{\mathbf{X}}_{\text{CTP}}^i$  to a binary mRS score and  $F_{\text{logistic}}$  is a mapping function of step 1 outputs and  $\mathbf{X}_{\text{Cli\_var}}^i$  to a binary mRS score.  $\hat{\mathbf{X}}_{\text{CTP}}^i$  is composed of  $\hat{\mathbf{X}}_{\text{CBF}}^i, \hat{\mathbf{X}}_{\text{CBF}}^i, \hat{\mathbf{X}}_{\text{MTT}}^i, \hat{\mathbf{X}}_{\text{Delay}}^i$ , concatenated as channels.

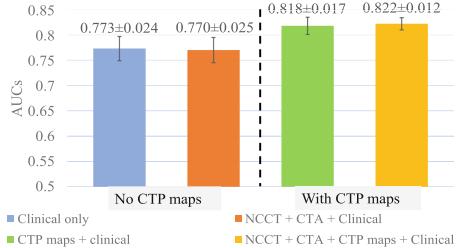
### 3 Experiments and Results

We performed two sets of experiments in the current study. In the first set of experiments, we compared prognostic prediction performance between models using different modalities, including (i) NCCT and CTA, (ii) CTP maps, (iii) NCCT, CTA and CTP maps, (iv) non-imaging data, (v) NCCT, CTA and non-imaging data, (vi) CTP maps and non-imaging data, and (vii) NCCT, CTA, CTP maps and non-imaging data. Images were input into the models with the architecture described in Sect. 2.2 stage 2, where inputs were replaced with corresponding imaging modalities concatenated at the channel level. In the second set of experiments, we evaluated the quality of the synthetic images and the performance when using them for prognostic prediction. We initially compared our model to four synthesis models: UNET, WGAN, CycleGAN and L2GAN. The L2GAN has the same architecture as our model but is not assigned the additional clinical task in the discriminator. To evaluate the quality of the synthetic images, we compared the structural similarity index measure (SSIM) and peak signal-to-noise ratio (PSNR) between synthesis models. Area under the ROC curve (AUC), accuracy (ACC), and F1-Score were used to assess the performance of prognostic prediction. We also compared our model to three state-of-the-art models that used raw images and clinical non-imaging data [4,5].

We randomly split the data into a training and a testing dataset. We used 4-fold cross-validation for training. For image synthesis, the models were trained for 200 epochs with a batch size of 4 using the Adam optimizer with learning rates of  $2 \times 10^{-4}$  and  $2 \times 10^{-5}$  for the generator and discriminator, respectively. For prognostic prediction, models with image inputs were trained for 100 epochs with a batch size of 4 using the Adam optimizer with a learning rate of  $1 \times 10^{-5}$ . All of the models were trained independently. The experiments above were implemented using PyTorch on the NVIDIA 3090 24GB GPU. Logistic regression models were trained with hyperparameters using grid search (Supplementary Table S1) based on Scikit-learn.



**Fig. 3.** AUCs of models using imaging data only



**Fig. 4.** AUCs of models incorporating imaging and clinical non-imaging data

**Table 1.** Image quality of five generative models

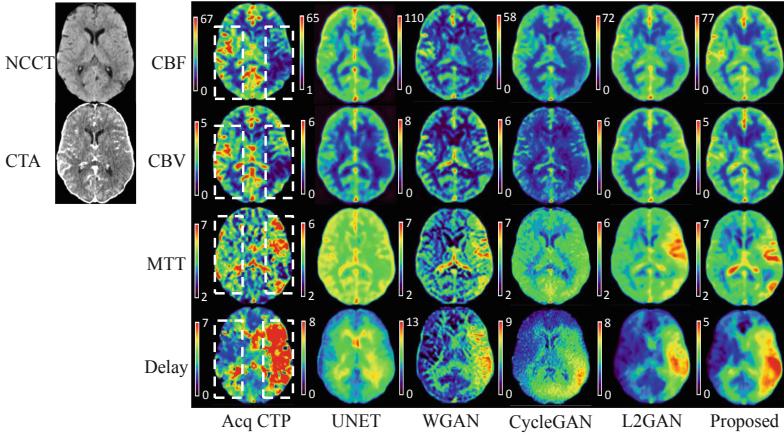
Models	SSIM (%)				PSNR			
	CBF	CBV	MTT	Delay	CBF	CBV	MTT	Delay
UNET	75.2 ± 0.5	79.3 ± 0.6	54.4 ± 1.3	61.7 ± 0.5	31.0 ± 0.2	34.7 ± 0.4	<b>20.0 ± 0.1</b>	<b>20.3 ± 0.1</b>
WGAN	79.0 ± 1.2	83.2 ± 0.8	46.4 ± 4.0	55.2 ± 2.1	31.0 ± 1.0	34.2 ± 0.5	19.5 ± 0.4	19.1 ± 0.3
CycleGAN	80.9 ± 1.2	84.5 ± 1.0	57.7 ± 0.5	66.6 ± 0.5	31.0 ± 0.3	34.1 ± 0.6	19.3 ± 0.1	19.3 ± 0.1
L2GAN	82.2 ± 0.3	85.8 ± 0.4	59.8 ± 0.5	67.7 ± 0.7	<b>31.7 ± 0.1</b>	35.1 ± 0.2	19.9 ± 0.3	20.2 ± 0.3
Ours	<b>82.5 ± 0.4</b>	<b>86.6 ± 0.2</b>	<b>60.7 ± 0.4</b>	<b>69.2 ± 0.5</b>	31.6 ± 0.3	<b>35.2 ± 0.1</b>	<b>20.0 ± 0.1</b>	20.2 ± 0.1

**Table 2.** Performance of models using acquired NCCT and CTA, and synthetic CTP maps, compared to the ideal case with acquired CTP maps (bottom row), for prediction of 3-month mRS.

	AUC(%)	ACC(%)	F1(%)
NCCT+CTA	77.0 ± 2.5	69.6 ± 3.9	68.0 ± 4.3
UNET	78.3 ± 1.0	69.3 ± 2.4	67.5 ± 3.0
WGAN	78.9 ± 3.3	70.9 ± 5.1	69.5 ± 6.0
CycleGAN	77.8 ± 1.9	70.1 ± 3.4	68.1 ± 3.4
L2GAN	79.5 ± 1.1	71.8 ± 2.0	69.5 ± 1.1
Our method	<b>80.7 ± 1.0</b>	<b>72.0 ± 2.4</b>	<b>70.3 ± 2.0</b>
Acquired CTP	81.8 ± 1.7	73.4 ± 4.2	71.4 ± 3.1

### 3.1 Results of Image Synthesis and Prognostic Prediction

**Data Modalities for Prognostic Prediction.** The performance of models using different combinations of data modalities is shown in Figs. 3 and 4 and supplementary Table S2. Models that included CTP maps clearly had better performance (Fig. 3). When non-imaging data were incorporated into the models (Fig. 4), those with CTP maps outperformed those without. This demonstrates that the inclusion of CTP maps can increase the performance of prognostic prediction. The validation results are similar to these test data results (Table S2).



**Fig. 5.** Illustration of synthetic CTP maps generated by models (rightmost 5 columns). The dashed squares indicate the location of the lesion in the acquired images (Acq CTP, which is the ground truth for image generation). The solid square indicates the non-ischemic regions for comparison. The colour bars show the corresponding CTP values of the colours.

**Table 3.** Performance of models for prediction of 3-month mRS. Bacchi S [4] and Samak Z [5] are state-of-the-art models. The bottom row shows the ideal case for comparison.

	AUC(%)	ACC(%)	F1(%)
Bacchi S <sub>NCCT only</sub> [4]	74.6 ± 3.0	69.6 ± 3.7	63.2 ± 5.9
Bacchi S <sub>NCCT+CTA</sub> [4]	76.8 ± 0.8	67.9 ± 0.6	63.4 ± 2.6
Samak Z [5]	76.6 ± 0.7	70.4 ± 1.4	63.7 ± 3.1
Our method	<b>80.7 ± 1.0</b>	<b>72.0 ± 2.4</b>	<b>70.3 ± 2.0</b>
Acquired CTP	81.8 ± 1.7	73.4 ± 4.2	71.4 ± 3.1

**Image Quality of Synthetic CTP Maps.** The synthetic CTP maps generated by different methods are shown in Fig. 5 and Supplementary Fig. S1. The acquired CBF and CBV maps show the ischemic core with hypointense signals on the right side, while MTT and Delay maps show the penumbra with hyperintense signals when compared to the contralateral side. Despite some visible differences from the acquired CTP maps, our synthesis method successfully learned the key clinical features related to the ischemic changes. The CTP maps generated by L2GAN are visually similar to our synthetic maps, except for the hypointense signals in the CBF and hyper-intense signals in Delay being weaker in the L2GAN. This shows that adding the clinical task enables the model to better learn key clinical features. Each of the other synthetic maps (UNET, WGAN and CycleGAN) missed even more of the important hyper- and hypointense fea-

tures across the synthesized CTP images. The proposed model also performed better than the other models in image quality metrics (SSIM and PSNR).

**Synthetic CTP Maps for Prognostic Prediction.** The model incorporating the CTP maps generated by our proposed method shows the best performance of prognostic prediction compared to other synthesis methods (Table 2). This indicates that the inclusion of another clinical task can improve the outcome prediction. Also, the predictive performance of our synthetic maps is considerably closer to that of the acquired maps (bottom row) (ROC curves shown in Supplementary Fig. S2), indicating that the proposed method can recover most of its predictive ability when CTP maps are unavailable. Moreover, our model with the synthetic CTP maps also outperformed three state-of-the-art models trained on our dataset (Table 3) (ROC curves shown in Supplementary Fig. S2). This demonstrates that training strategies incorporating CTP map synthesis may be able to encourage the models to concentrate more on the most clinically relevant features in NCCT and CTA images for outcome prediction. Such training strategies may help build models that not only have better performance but are also clinically trusted, given their ability to demonstrate the replication of key clinical imaging features. Validation set results are similar to these (Supplementary Tables S3–S5).

## 4 Conclusion

This study demonstrates that CTP maps, which are known to provide critical information for clinicians, also benefit prognostic prediction using deep learning techniques. When CTP maps are not available at hospital admission, their benefits can still be largely retained through image synthesis. Using multi-task learning with a simple clinical task, our model outperformed other synthesis methods in both image quality and the performance of prognostic prediction. Our synthetic CTP maps show key clinical features that are able to be readily discerned upon visual inspection. These findings verify the advantages of including additional CTP maps in LVO prognostication and establish the ability to effectively synthesize such maps to retain their benefits. While we acknowledge that our network architectures are not novel, we highlight the novelty of our architectures for stroke prognostication. The proposed framework can provide significant utility in the future to aid in the selection of patients for high-stakes time-critical EVT, particularly for those who have limited access to advanced imaging. Furthermore, by demonstrating the key clinical imaging features, our framework may improve confidence in building a clinically trusted model.

## References

1. Malhotra, K., Gornbein, J., Saver, J.L.: Ischemic strokes due to large-vessel occlusions contribute disproportionately to stroke-related dependence and death: a review. *Front. Neurol.* **8**, 651 (2017)

2. Powers, W.J., et al.: 2018 guidelines for the early management of patients with acute ischemic stroke: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke* **49**(3), e46–e99 (2018)
3. Wolman, D.N., et al.: Endovascular versus medical therapy for large-vessel anterior occlusive stroke presenting with mild symptoms. *Int. J. Stroke* **15**(3), 324–331 (2020)
4. Bacchi, S., Zerner, T., Oakden-Rayner, L., Kleinig, T., Patel, S., Jannes, J.: Deep learning in the prediction of ischaemic stroke thrombolysis functional outcomes: a pilot study. *Acad. Radiol.* **27**(2), e19–e23 (2020)
5. Samak, Z.A., Clatworthy, P., Mirmehdi, M.: Transop: transformer-based multimodal classification for stroke treatment outcome prediction. arXiv preprint [arXiv:2301.10829](https://arxiv.org/abs/2301.10829) (2023)
6. Zeng, M., et al.: Pre-thrombectomy prognostic prediction of large-vessel ischemic stroke using machine learning: a systematic review and meta-analysis. *Front. Neurol.* **13**, 945813 (2022)
7. Saleem, Y., et al.: Acute neurological deterioration in large vessel occlusions and mild symptoms managed medically. *Stroke* **51**(5), 1428–1434 (2020)
8. Nogueira, R.G., et al.: Thrombectomy 6 to 24 hours after stroke with a mismatch between deficit and infarct. *New Engl. J. Med.* **378**(1), 11–21 (2018)
9. Bal, S., et al.: Time dependence of reliability of noncontrast computed tomography in comparison to computed tomography angiography source image in acute ischemic stroke. *Int. J. Stroke* **10**(1), 55–60 (2015)
10. Chu, Y., et al.: Comparison of time consumption and success rate between CT angiography-and CT perfusion-based imaging assessment strategy for the patients with acute ischemic stroke. *BMC Med. Imaging* **22**(1), 1–8 (2022)
11. Chung, C.Y., Hu, R., Peterson, R.B., Allen, J.W.: Automated processing of head CT perfusion imaging for ischemic stroke triage: a practical guide to quality assurance and interpretation. *Am. J. Roentgenol.* **217**(6), 1401–1416 (2021)
12. Goodfellow, I., et al.: Generative adversarial networks. *Commun. ACM* **63**(11), 139–144 (2020)
13. Arjovsky, M., Chintala, S., Bottou, L.: Wasserstein generative adversarial networks. In: International Conference on Machine Learning, pp. 214–223. PMLR (2017)
14. Zhu, J.-Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2223–2232 (2017)
15. Jenkinson, M., Smith, S.: A global optimisation method for robust affine registration of brain images. *Med. Image Anal.* **5**(2), 143–156 (2001)
16. Muschelli, J.: A publicly available, high resolution, unbiased CT brain template. In: Lesot, M.-J., et al. (eds.) *IPMU 2020. CCIS*, vol. 1239, pp. 358–366. Springer, Cham (2020). [https://doi.org/10.1007/978-3-030-50153-2\\_27](https://doi.org/10.1007/978-3-030-50153-2_27)
17. Waaijer, A., et al.: Reproducibility of quantitative CT brain perfusion measurements in patients with symptomatic unilateral carotid artery stenosis. *Am. J. Neuroradiol.* **28**(5), 927–932 (2007)
18. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) *MICCAI 2015. LNCS*, vol. 9351, pp. 234–241. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)