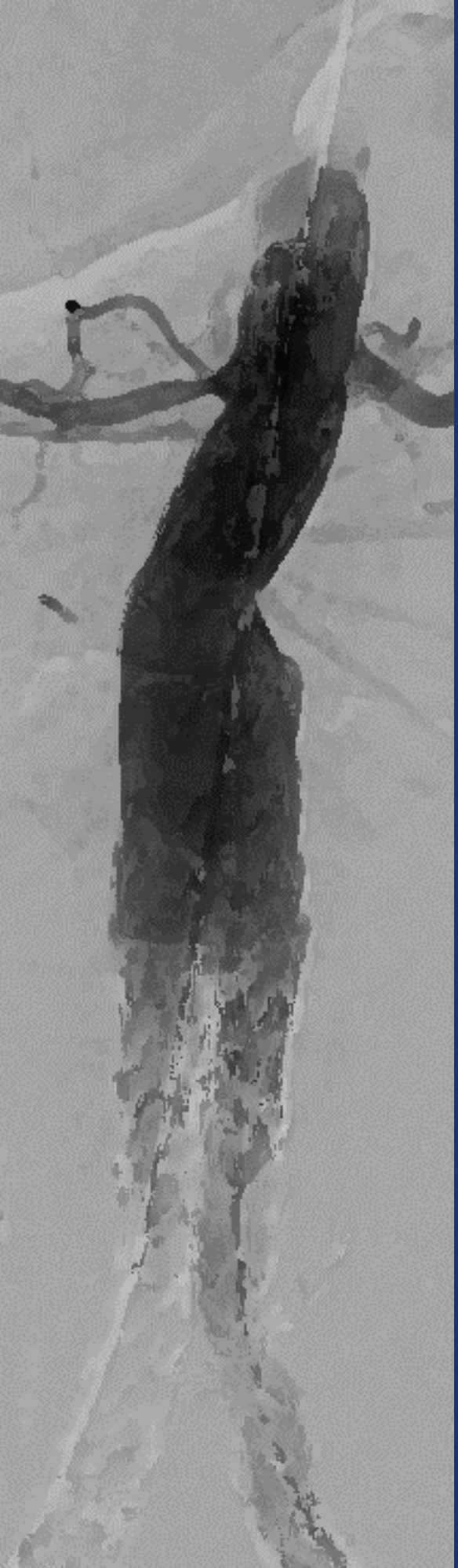

# Deep Learning for Objective Intraoperative Image Analysis during Endovascular Aneurysm Repair: *Automatic Artery Detection*

MSc. Thesis

Technical Medicine – Medical Imaging & Interventions

Kaj Olav Kappe

University of Twente
May 2022

UNIVERSITY OF TWENTE.

Amsterdam UMC
University Medical Centers

# Deep Learning for Objective Intraoperative Image Analysis during Endovascular Aneurysm Repair: *Automatic Artery Detection*

Master Thesis

*Technical Medicine – Medical Imaging and Interventions*

by

Kaj Olav Kappe

May 2022

Department of Vascular Surgery
Amsterdam University Medical Centers

# Graduation Committee

**Chairman**
Prof. dr. C. Brune
*Mathematics of Imaging & AI Group, University of Twente, Enschede*

**Medical Supervisor Institution**
Dr. K.K. Yeung, Vascular Surgeon
*Department of Vascular Surgery, Amsterdam University Medical Centers, Amsterdam*

**Technical Supervisor (UT)**
Dr. J.M. Wolterink
*Mathematics of Imaging & AI Group, University of Twente, Enschede*

**Technical Physician Supervisor Institution**
S.P.M. Smorenburg, MSc.
*Department of Vascular Surgery, Amsterdam University Medical Centers, Amsterdam*

**Process Supervisor**
E.M. Walter, MSc. / Drs. P.A. van Katwijk
*Technical Medicine, University of Twente, Enschede*

**External Member (UT)**
J.K. van Zandwijk, MSc.
*Magnetic Detection & Imaging, University of Twente, Enschede*

## Preface

After more than seven and a half year of studying, this thesis marks the end of my time as a Technical Medicine student. It has been an incredible time which I will look back at with great gratitude for the rest of my life. The past year, I was given the opportunity to fulfil my final clinical internship at the department of vascular surgery at the Amsterdam UMC. A time that boosted my development towards a Technical Physician in terms of research and clinical experience.

The past year would not have been possible with the help of many people. First of all, I would like to express my great appreciation to my direct supervisors of the last year. Kakkhee, your passion for vascular surgery and research is inexhaustible and you keep pushing me to strive for more. Your willingness to teach me the different aspects of vascular surgery in research, at the outpatient clinic and in the operating rooms are something I really appreciate. Jelmer, your expertise in artificial intelligence is inspiring and drives me to gain more knowledge in this domain. You have given me so many insightful ideas during our meetings and taught me to be less critical about my work at times. Stefan, you were always very approachable and it was always great to have our discussions from a Technical Physician's point of view. I look forward to them in the future. Furthermore, you have taught me how to fulfil a valuable role as a Technical Physician in the hybrid operating room. Elyse, you have pushed me to reflect on myself at a deeper level over the past two years. This has enabled me to develop myself as a Technical Physician but also on a personal level. Christoph, even though we haven't met often during the past year, I appreciate the discussions we have had on my research which led to new insights and interesting ideas.

Maarten, Lotte and Eva, we have spent many days together working on our research over the past year. I would like to thank you for the many interesting discussions we have had but also for the great atmosphere in the O2-building. Dieuwertje, you were always very willing to think along and we have shared many thoughts on both our research projects the past year and I hope we can keep sharing our knowledge in the coming years. Furthermore, I would like to thank all other researchers and vascular surgeons at the department of vascular surgery at the Amsterdam UMC for the great time over the past year.

At last, I want to express my great gratitude to my parents and Thomas for the amazing support that you always give me, whatever direction I go.

Enjoy reading this thesis!

Kaj Kappe
*Utrecht, 2022*

## Abstract

**Background:** Endovascular Aneurysm Repair (EVAR) is the predominant choice for elective and acute treatment of patients with an infrarenal abdominal aortic aneurysm. Completion digital subtraction angiography (cDSA) is performed at the end of an EVAR procedure to evaluate the position and patency of the stent-graft, potential endoleaks and blood flow dynamics. Clinical decision-making is based on the cDSA. However, the evaluation is based on visual inspection by the surgical team and therefore subjective and prone to inter-observer disagreement. In this study, we aim to use deep learning to automatically extract imaging-based features, in particular detection of renal and lumbar arteries, as part of the development of objective intraoperative image analysis.

**Methods:** In chapter 2, we investigated the visibility of lumbar arteries between the pre-operative Computed Tomography Angiography (CTA) and cDSA in thirteen patients treated with EVAR.

In chapter 3, a U-Net was trained for the automatic detection of renal arteries on the cDSA, based on two-dimensional (2D) feature projections obtained from the cDSA. The performance of the model was evaluated using the median localization error and successful cumulative detection rate at different distance thresholds and for different combinations of input feature projections.

In chapter 4, we developed a deep-learning based method, using U-Net, for the automatic identification of lumbar arteries on the cDSA, using 2D feature projections extracted from the cDSA. The area under the curve of the precision-recall curve ($AUC_{PR}$) was used to evaluate the performance of the model for different combinations of input feature projections.

**Results:** Over-projection of lumbar arteries by the aortic lumen on the cDSA occurred in 2% of the cases. Invisibility of the lumbar artery on the pre-operative CTA is a good predictor for invisibility on the cDSA, while visibility on the pre-operative CTA was a moderate predictor for visibility on the cDSA.

The median localization error of renal arteries for the different input projection combinations on the cDSA were 1.52, 1.52, and 1.65 mm (p=0.853). The successful cumulative detection rate was around 75% for all input projection combinations.

The $AUC_{PR}$ was 0.59, 0.61, 0.62 and 0.67 for the identification of lumbar arteries on the cDSA for the different combinations of input feature projections.

**Conclusion:** This study has demonstrated the feasibility of using deep learning models for the automatic detection and identification of renal and lumbar arteries on the cDSA. These extracted features can serve as building blocks for the development of objective intraoperative image analysis during EVAR.

# Contents

Chapter 1

# Introduction to Abdominal Aortic Aneurysms and Endovascular Aneurysm Repair

## 1.1 Abdominal Aortic Aneurysm

An Abdominal Aortic Aneurysm (AAA) is an irreversible localized dilation of the abdominal aorta.[1] The most common site for an AAA to develop is below the renal arteries, and is called an infrarenal AAA.[1] A dilation of the infrarenal aorta is called an AAA when the diameter of the vessel is 1.5 times the normal diameter or when the diameter is ≥ 30 millimetres.[1–3] The most important risk factors for the development of an AAA are smoking, male sex, age and family history.[1] AAA's are most commonly diagnosed incidentally during examination of another pathology as patients with an AAA are usually asymptomatic.[1] According to the European Society for Vascular Surgery (ESVS) 2019 Clinical Practice Guidelines, elective repair of the AAA should be considered when the diameter equals or exceeds 5.5 cm for men and the threshold for elective repair is 5.0 cm for women.[2] If the aneurysm is left untreated, the risk of rupture grows with an increasing diameter with a mortality rate of 70-80% for ruptured AAA's.[4,5] Elective surgical repair of an AAA can be performed with invasive open surgery in which a vascular prosthesis replaces the aneurysm or with minimal invasive surgery in which a stent-graft is endovascularly implanted into the aneurysm, as shown in Fig. 1.[1]

## 1.2 Endovascular Aneurysm Repair

Endovascular aneurysm repair (EVAR) has increasingly established itself as the predominant choice for elective and acute treatment for patients with an infrarenal AAA, as an alternative to the more invasive traditional open repair. [6] In recent years, around 75% of the patients undergoing elective aneurysm repair were treated using EVAR in The Netherlands.[6] During a minimally invasive EVAR procedure, intraluminal access to the aortic aneurysm is realized through a surgical cut down or percutaneous puncture to the femoral arteries.[1,7] With the use of catheters, guide-wires and stent-graft introducer systems, the bifurcated endoprosthesis is placed and deployed inside the aneurysm sac below the renal arteries under guidance of intraoperative C-arm fluoroscopy imaging.[1,7,8] Proximal sealing of the stent-graft is achieved by anchoring the stent in the non-aneurysmal aortic wall below the renal arteries using radial force of the nitinol stent and inflation balloon catheter.[1] The distal end of the stent-graft is sealed in the common iliac artery in a similar fashion.[1] After deployment and an adequate proximal and distal seal of the stent-graft, the AAA is excluded from the systemic blood flow while leaving the aneurysm sac in situ rather than replacement of the dilated aorta by a prothesis as in the traditional open surgery.[1] The aim of aneurysmal sac exclusion from the systemic circulation is the absence of continuous pressure on the dilated aortic wall to prevent further growth and possible future rupture.[9]

## 1.3 Advantages and disadvantages of EVAR

Several randomized controlled trials have been conducted to compare the outcomes of EVAR with open AAA repair. The Dutch Randomized Endovascular Aneurysm Management (DREAM) trial and the United Kingdom EVAR-1 trial have shown the benefit of EVAR compared to open AAA repair in terms

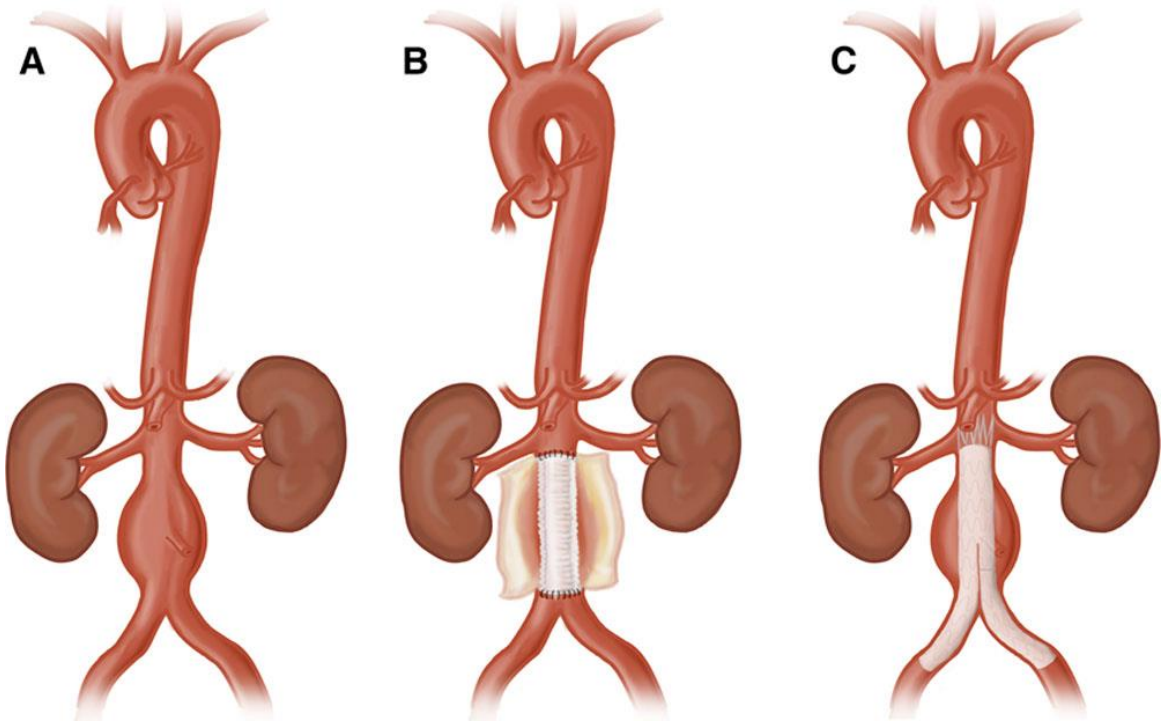


Fig. 1. A: Infrarenal abdominal aortic aneurysm. B: Open repair of an AAA with a vascular prosthesis. C: Endovascular aneurysm repair with an inserted stent graft through the femoral arteries. Image from Swerdlow, et al.[10]

of 30-day mortality.[8,11] Furthermore, endovascular repair demonstrated significant shorter duration of the surgical procedure, less blood loss and a shorter stay on the intensive care and medium care unit as well as shorter total hospitalization.[8] The advantage of EVAR in terms of short-term survival, however, disappeared within two years after surgery and long-term follow-up of 12-15 years showed no significant difference in overall survival between patients treated with endovascular and open repair.[12,13]

Despite its advantage in short-term survival and similar long-term overall survival compared to open repair, EVAR is associated with a significant higher number of secondary interventions during follow-up.[12–14] Reintervention rates were 16.7 percentage points higher for patients treated with EVAR after twelve years of randomization.[13] Endoleaks are the most commonly encountered complications following EVAR, followed by other endograft-related complications such as stent-graft migrations and endograft limb occlusion and these are all indications for reinterventions.[9,12,14,15] Endoleaks are defined as persistent blood flow into the aneurysm sac after exclusion from the systemic circulation by stent-graft deployment.[9,15]

### 1.4 Types of Endoleaks

Endoleaks can be classified into five types, according to the underlying origin of the inflow of blood into the aneurysm sac, as illustrated in Fig. 2.[9,16] Endoleaks of any type occur in about one third of the patients treated with EVAR.[17] Type I endoleaks result from an incomplete circumferential seal and apposition

between the stent-graft and the arterial wall. Type I endoleaks can be subdivided according to the sealing zone location, either at the proximal end of the stent-graft (type Ia) or the distal end in the iliac arteries (type Ib).[15,18] The insufficient seal leads to direct aortic inflow of blood into the peri graft cavity and exposes the aneurysm sac to high systemic pressures, requiring urgent treatment to prevent aneurysm sac expansion and possible rupture.[18–20] Type I endoleaks as a result of suboptimal device selection and incomplete dilation of the stent-graft, which has insufficient radial force into the native aortic wall, are usually visible intraoperatively[15,20] Furthermore, challenging anatomy such as severe supra-renal neck angulations, existence of intraluminal thrombus and a short sealing zone in the non-aneurysmal proximal aortic neck are associated with a higher risk of type I endoleak development.[21] Later onset of type I endoleaks may occur after alteration in the aortic morphology after disease progression, aneurysm sac growth due to a type II endoleak or aneurysm sac shrinkage.[15,20,22] Several endovascular options exist for treatment of type 1 endoleaks, such as further balloon-dilation of the stent-graft, placement of a bare metal stent or the use of endovascular staples to achieve apposition of the stent-graft fabric to the aortic wall.[2]

Type II endoleaks are a retrograde flow of blood from an excluded aortic side branch into the aneurysm sac after deployment of the stent-graft.[20] Typical sources of retrograde flow into the aneurysm sac are the inferior mesenteric artery, lumbar arteries and the median sacral artery.[20,23] Simple type II endoleaks consist of a single aortic branch as the source, complex type II endoleaks include multiple involved aortic side branches.[23] Type II endoleaks are the most common type of endoleak and occur in 20-25% of the patients and make up around 75% of all endoleaks detected in a sub-analysis of the patient cohort that participated in the OVER-trial.[17,24] Type II endoleaks can further be classified into early- and late-onset in which the early-onset endoleaks can be subclassified into transient and persistent endoleaks. Transient endoleaks spontaneously thrombose within 6 months after treatment and persistent endoleaks last over 6 months.[25] Persistent type II endoleaks may lead to aneurysm sac expansion over time, however, there is an ongoing scientific debate on the best clinical and therapeutic management on patients with persistent endoleaks that lead to sac enlargement.[26] The European Society for Vascular Surgery (ESVS) recommends reintervention if the expansion of the aneurysm sac diameter exceeds 10 millimetres during follow-up.[2] Common treatment of type 2 endoleaks are transarterial and translumbar embolization of the aneurysm sac and arteries that cause the endoleak.[2]

Type III endoleaks occur after disconnection of the stent-graft components or a defect of the graft's covering fabric material and may have an early or late onset.[20,27] Deployment of stent-graft components with insufficient overlap, migration and material fatigue are the underlying mechanism of type III endoleaks.[2,27] The estimated frequency of a type III endoleak in long-term follow-up is 1-3%, however, these numbers also include older generation stent-grafts and type III endoleaks are less frequently seen in newer generation stent-grafts.[2,27] Urgent management of type III endoleaks, primarily by endovascular means using relining of the stent-graft, is required due to its high pressure nature and common association with aneurysm sac expansion and possible rupture.[2,20]

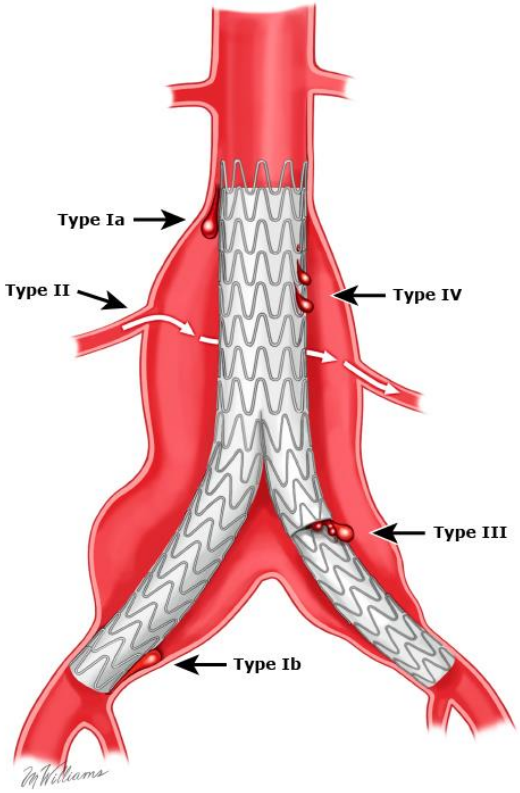


Fig. 2. Illustration of different types of endoleaks. Image from of UpToDate 2022.[28]

Type IV endoleaks are defined as leakage of blood into the aneurysmal sac due to the porosity of the stent-graft's covering fabric.[2] Advancements in stent-graft materials have made this type of endoleak very rare and does not require reintervention.[2,29] Type V endoleaks are described as aneurysm sac expansion without direct evidence of leakage, also described as endotension.[20]

## 1.5 Management of type 2 Endoleaks

The clinical significance and therapeutic management of type 2 endoleaks remains controversial as their natural aetiology is still a poorly understood topic.[26,30] Part of the type 2 endoleaks resolve spontaneously, however, persistent type 2 endoleaks have been associated with aneurysm sac enlargement, need for reintervention and conversion to open repair.[25,30] A study by Lo et al. demonstrated a significantly higher rate of reinventions in patients with a persistent type 2 endoleak compared to patients without an endoleak or a transient endoleak, but there was no difference between the groups in terms of rupture and survival.[31] Type 2 endoleaks are however associated with an increased risk of aneurysm-related mortality.[26] The current ESVS guidelines recommend reintervention for type 2 endoleaks in case of aneurysm sac diameter expansion exceeding 10 millimetres during follow-up.[2]

As mentioned previously, aneurysm sac expansion is associated with a higher rate of reintervention.[31] In contrast, aneurysm shrinkage is considered a sign of true success of aneurysm sac exclusion from the

systemic blood circulation after EVAR. A meta-analysis by Antoniou et al. demonstrated that patients with aneurysm sac shrinkage have significantly better outcomes than patients with a stable or expanding aneurysm sac following EVAR.[32] Furthermore, patients with a shrinking AAA one year after EVAR have a better long-term survival.[33] Early prognostic knowledge on the risk of sac expansion or shrinkage based on imaging-based parameters may help classify treated patients into a high and low-risk group. This can lead to a personalized follow-up and possibly earlier treatment to prevent adverse outcomes.

Several studies have been conducted to investigate possible imaging-based predictive factors associated with aneurysm sac behaviour after diagnosis of a type 2 endoleak. Azofra et al. investigated imaging-based features on the first post-operative CTA for the prediction of aneurysm sac enlargement. They found that the craniocaudal length and the nidus volume of the endoleak were significantly associated with aneurysm sac growth.[26] These results are in line with the results of Demehri et al. who concluded that the endoleak cavity volume was a predictive factor for aneurysm sac growth.[9] The complexity of the type II endoleak and the diameter of the largest connected artery to the aneurysm sac on the first post-operative CTA within 30 days were predictive indicators for aneurysm sac growth in a study by Müller-Wille et al.[23]

The number of patent lumbar arteries on the pre-operative CTA has been identified as a risk factor for the occurrence of a type 2 endoleak after EVAR but also as a predictive factor for the persistence of a type II endoleak after EVAR.[34,35] Besides the relationship between the persistence of a type II endoleak with the number of patent lumbar arteries, Seike et al. also demonstrated a significant relationship between aneurysm sac enlargement and the number of patent lumbar arteries.[35] In their study, they report that patients with four or more lumbar arteries on the pre-operative CTA should be considered as a high-risk group for sac enlargement after EVAR.[35]

### 1.6 Workflow in the hybrid operating room

In the Amsterdam University Medical Centers, patients with an infrarenal AAA are treated in a Philips Azurion FlexMove 7 C20 Hybrid Operating Room (Philips Healthcare, Best, The Netherlands). During the procedure, fluoroscopic imaging is used for navigational purposes and correct placement and precise deployment of the endoprosthesis inside the aneurysm sac. Digital subtraction angiography (DSA) is a fluoroscopic imaging technique used for the visualization of blood vessels and plays an important role during the EVAR procedure. In this technique, a series of images is recorded during the injection of a contrast-agent. Subsequently, a pre-contrast-injection image is digitally subtracted from the post-injection series resulting in a series of images solely showing the contrast-filled blood vessels. DSA provides high-quality images with high temporal and spatial resolutions and the ability to represent the dynamics of the injected contrast.[36,37]

Fig. 3. Images from a completion DSA series recorded at the end of an EVAR procedure. The red arrows indicate the renal arteries, the orange arrows indicate the lumbar arteries, and the orange box outlines a type 2 endoleak.

DSA is utilized during the procedure to precisely place and deploy the stent-graft below the renal arteries and proximal to the internal iliac artery. Furthermore, DSA is used to assess the stent-graft position, ensure patency of the stent-graft and arteries and it also remains the gold standard for the detection and classification of endoleaks at the end of the EVAR procedure by means of a completion angiogram.[2,36] During a completion angiogram, also known as a completion DSA (cDSA), contrast-agent is injected through a diagnostic catheter such as a pigtail, which is placed proximally to the top of the stent-graft and renal arteries. The obtained images are visually inspected by the surgical team on the aforementioned aspects. An example of a cDSA series is shown in Fig. 3.

In a study performed by Smorenburg et al. (to be published), the final completion angiograms of 102 EVAR procedures were retrospectively analysed by two vascular surgeons and two interventional radiologists. The vascular surgeons and interventional radiologists independently scored the final completion angiogram on the presence and classification of endoleaks. The scoring of the vascular surgeons, interventional radiologists and the surgical team that performed the procedure were individually compared with each other. Results of the individual comparisons showed an interobserver agreement's

Cohen's Kappa score ranging from 0.302 to 0.561 on the detection of endoleaks on the final completion angiograms, excluding possible disagreement on the classification if two specialists agreed on the existence of an endoleak. From the study by Smorenburg et al., we can conclude that detection and classification of endoleaks on a completion angiogram is far from trivial and has large variability among the observers.

## 1.7 Rationale for the current study

Clinical decision-making during EVAR procedures is dependent on the subjective visual inspection of the completion angiogram by the surgical team. As Smorenburg et al. has demonstrated, this subjective evaluation can lead to a reduced inter-observer agreement, potentially influencing the decision-making and outcome of the procedure. While this study only evaluated the inter-observer agreement with regards to endoleaks, other aspects such as stent-graft position are also evaluated by visual inspection and thus prone to individual interpretation by the surgical team members as well.

Automatic objective analysis of the completion angiogram can help to detect regions of interest that need further inspection and reduce the inter-observer variability amongst the surgical team members. Early and intraoperative objective evaluation of the completion angiogram can aid in management of possible encountered problems within the same procedure, therefore preventing a potential reintervention.

The vascular research group in the Amsterdam University Medical Centers focuses on the automatic intraoperative evaluation of the completion angiogram using deep learning. Deep learning is able to objectively analyse the intraoperative images based on previously acquired data and can assist the surgical team during decision-making and reduce inter-observer disagreement. Furthermore, the use of deep learning for the analysis has no additional burden to the surgical team during the intervention.

This thesis focuses on the automatic detection of renal arteries and lumbar arteries. The automatic detection of these arteries can aid in several objective analytical tools for the completion angiogram. The automatic detection of the renal arteries can be used for stent-graft deployment evaluation and classification of transient and persistent type 2 endoleaks, according to Mursalin et al.[25] The detection and localization of lumbar arteries can furthermore be used for classification of type 2 endoleaks, as these arteries, along with the inferior mesenteric artery, account for the majority of the type 2 endoleaks.[38]

The overall aim of this study is to automatically extract imaging-based features, focussing on arteries, from the completion angiogram with a convolutional neural network that can be used for the objective intraoperative analyse of the final completion angiogram at the end of an EVAR procedure. In chapter 2, the objective is to compare the visibility of lumbar arteries on the pre-operative CTA and completion angiogram and to investigate a possible correlation between the visibility of lumbar arteries between the two imaging modalities. In chapter 3 and 4, the objective is to develop deep-learning based methods for automatic detection of renal arteries and identification of lumbar arteries on final completion angiogram.

Chapter 2

# Lumbar artery visibility correlation between preoperative computed tomography angiography and completion digital subtraction angiography: a qualitative analysis

## 2.1 Introduction

Endovascular aneurysm repair (EVAR) has beneficial peri-operative outcomes for the treatment of infrarenal abdominal aortic aneurysms compared to conventional open surgical repair.[8,39] Despite these early benefits, the endovascular surgical approach shows more aneurysm-related complications with an indication for reintervention, such as endoleaks.[40] Type 2 endoleaks are the most frequently encountered type of endoleak and are a result of retrograde flow of blood into the aneurysm sac from the inferior mesenteric artery and lumbar arteries.[17] Ward et al. has demonstrated that an increased number of patent lumbar arteries on the pre-operative CTA is associated with an increased risk of a type 2 endoleak.[38]

Pre-operative planning of an EVAR procedure and inspection of the patient's anatomy is performed on 3-dimensional Computed Tomography Angiography (CTA) scans. CTA imaging enables the vascular surgeon to view the patient's anatomy, in particular contrast-filled arteries, in different planes. These different planes allow the surgeon to pre-operatively examine the presence and course of arteries rather than in a single plane. Preoperatively, CTA imaging is also analysed in terms of the number of lumbar arteries as this is associated with an increased risk of a type 2 endoleak.[38]

In contrast to pre-operative CTA imaging, intraoperative imaging in the hybrid operating room is predominantly performed in 2D, thus visualizing the patient's vascular anatomy in a single projection plane. In comparison with multiplane imaging, single plane imaging is prone to over-projection of contrast-filled arteries making it more difficult to visualize arteries and distinguish them from each other. This also counts for the acquisition of the completion digital subtraction angiography (cDSA) at the end of the procedure which is used to inspect the stent-graft's position, patency of the stent-graft's lumen and also to check for possible endoleaks. This evaluation, however, is still performed by visual inspection and thus prone to interobserver variability.

At the Amsterdam University Medical Centers, research focuses on the development of automatic objective intraoperative analysis tools for the evaluation of the completion angiograms using deep learning, such as detection and classification of endoleaks. As part of this research, we also aim to develop a deep-learning based method for the automatic detection of lumbar arteries on the completion angiogram. Due to the correlation between the number of lumbar arteries and type 2 endoleaks, the automatic detection of these arteries in proximity of an endoleaks on the completion angiogram can aid in the automatic classification of endoleaks.

To date, no previous studies have investigated a possible correlation between the visibility of the lumbar arteries on the pre-operative CTA and completion angiogram. For the reliability of an automatic method for detection of lumbar arteries, it is important to investigate the visibility of these lumbar arteries on the completion angiogram as they need to be visible in the single projection plane to be detected. However, the single projection plane acquisition of the completion angiogram can lead to reduced visibility of the lumbar arteries due to over-projection. This can result in reduced detection of patent lumbar arteries by

the automatic deep learning-based method and therefore limit the reliability of eventual endoleak classification.

In this retrospective sub-analysis, we sought to compare the visibility, and investigate whether there is a correlation, between the lumbar arteries on the pre-operative CTA and the intraoperative cDSA images of patients treated with EVAR. We hypothesize that over-projection may lead to a reduced number of visible lumbar arteries detected on the cDSA compared to the pre-operative CTA.

## 2.2 Methods

In this retrospective sub-analysis, thirteen patients with an infrarenal AAA were included. The included patients were treated using EVAR in a Philips Azurion FlexMove 7 C20 Hybrid Operating Room (Philips Healthcare, Best, The Netherlands) in the Amsterdam University Medical Centers, location VUmc, between April 2017 and August 2019. The average aneurysm diameter was 61 ± 9 millimetre and all patients received a bifurcated endoprosthesis. Two patients additionally received an iliac branch device for an iliac aneurysm. The pre-operative CTA and cDSA imaging were retrospectively collected from a prospectively maintained database. Analyses of the images was performed in RadiAnt, a freely available PACS DICOM viewer for medical images (Medixant, Poznan, Poland).

First, the pre-operative CTA of all patients was analyzed for the visibility of lumbar arteries on the vertebral levels L1 to L5. To this end, the pre-operative CTA was evaluated in the axial, saggital and coronal planes to identify and follow the course of the lumbar arteries. The different planes for two lumbar arteries are illustrated in Fig. 4. The lumbar arteries on the pre-operative CTA were scored as visible reaching the aorta (visible), visible but not reaching the aorta (partly visible) or invisible by a single examiner. The lumbar arteries that were scored as partly visible were distally visible but were not visible near the origin at the aorta due to a lack of contrast-agent.

Secondly, a maximum intensity projection (MaxIP) of the pre-operative CTA was made along the coronal direction to reconstruct a partly comparable image of the expected image on the cDSA during surgery. The MaxIP of the CTA is used to identify the location and course of the lumbar artery in the same coronal projection on the cDSA, as shown in Fig. 5. The renal arteries on the cDSA and the MaxIP are used as a reference for identification of the lumbar arteries at the same vertebral level on both image modalities.

Using the set-up as demonstrated in Fig. 5, the lumbar arteries on the cDSA were scored as visible, over-projected by contrast-agent in the aortic lumen, undistinguishable from other contrast-agent filled blood vessels at the expected location of a lumbar artery or invisible. In the analysis, the pre-operative CTA was considered as the gold standard, and the visibility of the lumbar arteries on the cDSA was compared to the pre-operative CTA.

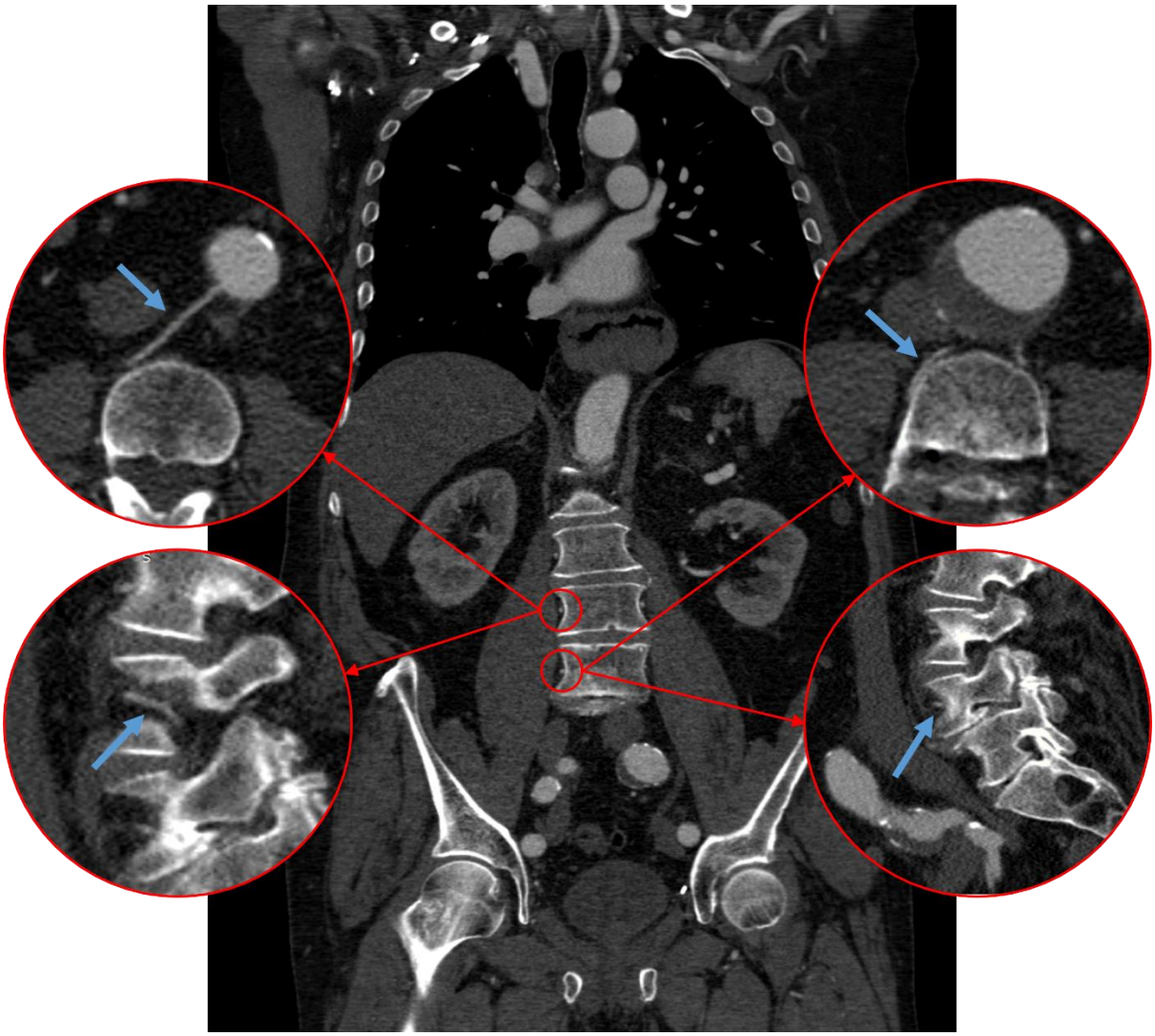

Fig. 4. Illustration of lumbar arteries indicated by the blue arrows at the vertebral levels L3 and L4 viewed on the coronal plane (large image), axial plane (upper circles) and sagittal plane (lower circles).

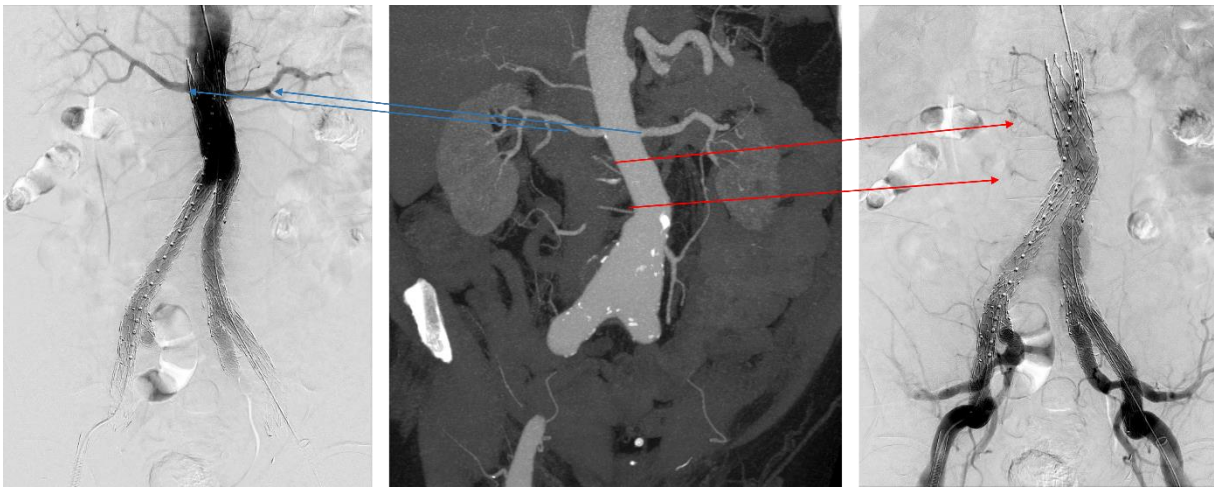

Fig. 5. Stills from the cDSA (left and right image) and the reconstructed MaxIP of the pre-operative CTA (middle) used for the identification of the lumbar artery's course and vertebral level. The red arrows indicate two lumbar arteries at levels L2 and L3. The blue arrows indicate the renal arteries that were used to determine the vertebral levels.

| | | Completion DSA | | | | |
|---|---|---|---|---|---|---|
| | | Visible | Undistinguishable* | Over-projected | Invisible | **Total** |
| **Pre-operative CTA** | Visible | 55 | 13 | 2 | 26 | **96** |
| | Partly visible | 0 | 2 | 0 | 2 | **4** |
| | Invisible | 0 | 3 | 0 | 27 | **30** |
| | **Total** | **55** | **18** | **2** | **55** | **130** |

Table 1. Scoring of the lumbar arteries on the cDSA compared to the pre-operative CTA. The pre-operative CTA is considered as the gold standard for presence of a lumbar artery. *Lumbar arteries were scored as undistinguishable if there was contrast-agent visible at the location of an expected lumbar artery but clear contours of the artery could not be identified.

## 2.3 Results

A total of 96 lumbar arteries were scored as visible on the pre-operative CTA image, from a total of 130 lumbar arteries that could theoretically be detected (10 lumbar arteries per patient, 13 patients). Table 1 shows the scoring of the lumbar arteries on the cDSA compared to the pre-operative CTA. If a lumbar artery was scored visible on the pre-operative CTA, the same lumbar artery was visible on the cDSA in 57% (55/96) of the cases and invisible in 27% (26/96). 14% (13/96) of the lumbar arteries scored as visible on the pre-operative CTA could not be clearly identified as lumbar arteries on the cDSA meaning that contrast-filled arteries were visible on the expected lumbar artery's location but a clear contour of the lumbar artery could not be defined. Two of the 96 (2%) of the lumbar arteries visible on the pre-operative CTA were over projected by the aortic lumen on the cDSA.

In contrast, if a lumbar artery was not visible on the pre-operative CTA, the same lumbar artery was also not visible on the cDSA in 90% (27/30) of the cases and was undistinguishable as a lumbar artery from other arteries in 3/30 (10%) of the cases. For the lumbar arteries that were scored as partly visible on the pre-operative CTA, two (50%) were scored as undistinguishable and two (50%) were scored as invisible on the cDSA.

Overall, the visibility of a lumbar artery on the pre-operative CTA has a moderate predictive value for the visibility of the same lumbar artery on the cDSA. However, if the lumbar artery is not visible on the pre-operative CTA, it has a good predictive value for the same lumbar artery to be invisible on the cDSA. Furthermore, there were only two cases of complete over projection of the lumbar artery by the contrast-filled aortic lumen.

## 2.4 Discussion

In this sub-analysis, we have demonstrated the correlation between the visibility or invisibility of a lumbar artery on the cDSA compared with the pre-operative CTA. The most valuable findings are that over-projection only occurs in a very few cases and that invisibility of the lumbar artery on the pre-operative CTA has a good predictive value for the invisibility of the same lumbar artery on the cDSA while the

visibility of a lumbar artery on the pre-operative CTA has a moderate predictive value for the visibility of the same lumbar artery on the cDSA.

The large number of invisible lumbar arteries on the cDSA compared to the pre-operative CTA could partly be explained by the presence of the stent-graft in the aorta at the time of the cDSA acquisition. In a healthy situation, the lumbar artery is supplied with blood directly from the aorta. During the pre-operative CTA, a quick and direct inflow of the contrast-agent into the lumbar arteries from the aorta is possible. The deployment of a stent-graft removes a direct connection between the aorta and the lumbar arteries, inhibiting fast and direct inflow of contrast-agent into the lumbar arteries from the aorta. In case of a placed stent-graft, the visibility of the contrast-filled lumbar arteries on the cDSA depends on collateral connections, for example through the hypogastric arteries or uncovered proximal lumbar arteries and intercostal arteries. The time before the contrast-agent reaches the lumbar arteries through collateral connections may exceed the recording time of the cDSA, therefore leading to invisiblity on the cDSA. Additionally, the collaterals also need to be sufficiently patent to allow contrast-agent inflow.

Another reason for the invisibility of the lumbar arteries on cDSA compared to the pre-operative CTA is that the contrast-agent may be diluted before reaching the lumbar arteries through the collaterals. Furthermore, the placement of the pigtail catheter and the settings for contrast injection on the cDSA also influences the flow of contrast-agent, while a reduced heart function may be another underlying cause for insufficient inflow through collaterals towards the lumbar arteries.

The invisibility of the lumbar artery on the pre-operative CTA seems to be a good predictor for the invisibility of the same lumbar artery on the cDSA. The exact correlation remains speculative, but may be due to the occlusion of a lumbar artery due to thrombosis at the time of the pre-operative CTA which does not resolve before the intervention. Therefore, the lumbar artery will not be visible on the cDSA.

Information obtained from this sub-analysis can be useful for the development of a deep learning-based method for the automatic detection and classification of endoleaks on a cDSA. An important finding in this sub-analysis regarding the development of the model is that complete over projection of the lumbar artery only occurs in a few percent of the cases. This reduces the chance of missed lumbar arteries by the model as a result of over projection. The good predictive value for the invisibility of a lumbar artery on the cDSA based on the invisibility on pre-operative CTA could be used as a form of prior information. Incorporation of this information may be of value for the automatic classification of endoleaks.

## 2.5 Conclusion

In conclusion, over-projection of the lumbar arteries on the cDSA occurs in a few percent of the cases based on this analysis. Furthermore, the pre-operative CTA can be considered as a valuable predictor of the invisibility of a lumbar artery on the cDSA. Information regarding the visibility of lumbar arteries on pre-operative CTA could be of added value in a deep learning based model for the automatic classification of endoleaks.

Chapter 3

# Deep learning-based detection of renal artery ostia on completion digital subtraction angiographies during Endovascular Aneurysm Repair

## 3.1 Introduction

Endovascular Aneurysm Repair (EVAR) is the preferred treatment of choice for infrarenal abdominal aortic aneurysms (AAA). In the Netherlands, nearly 80% of all patients that are treated electively for an intact infrarenal AAA and 40% of patients with a ruptured AAA undergo EVAR.[41] Compared to the conventional open surgical repair, the endovascular approach has proven its benefit in the early post-operative phase with less peri-operative mortality, reduced severe complications, less blood loss, and a shorter stay in the intensive care unit and total hospital stay.[8] On the other side, EVAR is significantly related to a higher rate of reintervention.[12]

The high rate of secondary interventions after initial treatment with EVAR remains the Achilles heel of the endovascular approach and endoleaks account for the majority of the complications that arise after placement of the stent-graft.[12] Acquisition of the completion angiogram, using digital subtraction angiography, is done at the end of an EVAR procedure to evaluate the position and patency of the stent-graft and visualize endoleaks. Clinical decision-making on whether or not further surgical treatment is warranted is based on visual inspection of the completion angiogram.

Even though the completion angiogram plays a crucial role in clinical decision-making, the analysis of the angiogram is performed by visual inspection by the surgeon and thus subjective. Very little research has been performed comparing the decision-making of individual surgeons, despite a study by Smorenburg et al. who have analysed the inter-observer agreement for the detection of endoleaks among surgeons. The moderate to fair inter-observer agreement in this study suggests room for the introduction of objective analytical tools for endoleak detection, but potentially also for other aspects such as stent-graft deployment accuracy.

In light of the transition towards an objective analysis of the completion angiogram, we propose to use deep learning techniques. In this study, we propose to automatically localize the renal artery ostia on the images of the completion angiogram. Renal artery ostia can be used in different manners for the objective analysis of the completion angiograms, as it is an important landmark for stent-graft deployment and may therefore aid in objective analysis of the stent-graft position. Furthermore, a study by Mursalin et al. has demonstrated that the appearance time of the endoleak cavity in relation to the appearance of the renal artery on the completion angiogram is a predictor of persistence for early on-set type 2 endoleaks.[25] Lastly, the renal artery ostia appearance can serve as a standardized reference point for further analysis of the completion angiogram. The aim of this study is to automatically localize the renal artery ostia using a deep learning model and investigate the influence of additional input channels on the performance of the network. The hypothesis is that the deep learning network is capable of correctly detecting the renal artery ostia on the cDSA and that additional input channels increase the performance of the network.

## 3.2 Methods

### *3.2.1 Proposed pipeline*

In this study, we aim to automatically detect the origin of the renal arteries on a completion digital subtraction angiography acquired at the end of an EVAR procedure. For localization of renal artery ostia, we optimize a Convolutional Neural Network with a U-Net[42] architecture that takes 2D projections of the cDSA series as input data and returns a prediction heatmap that represents the proximity of a given pixel to the renal artery origin's location. Subsequently, the location of each renal artery origin is found by extracting the pixel coordinates of the two highest activation sites on the prediction heatmap returned by the network.[43] In this study, we also investigate the effect of using a single 2D MinIP projection or multiple 2D projections obtained from the 2D + time cDSA as input on the performance of the network.

### *3.2.2 Data*

To develop our methods, we used a data set consisting of completion digital subtraction angiographies (cDSA) of patients that were treated for an infrarenal abdominal aortic aneurysm using EVAR. All cDSAs were acquired using a Philips Azurion FlexMove 7 C20 hybrid operating room (Philips Healthcare, Best, The Netherlands) in the Amsterdam University Medical Centers, location VUmc, between April 2017 and April 2020. Images were retrospectively extracted from a prospectively maintained database with the approval of the local medical ethical review committee.

A multiphase acquisition protocol was used for the acquisition of the cDSAs: during the first phase, images are acquired at a frame rate of 3 frames per second (fps) for 6 seconds, during the second phase, images are acquired at 1 fps for 5 seconds and images are acquired at 0.5 fps during the third phase for the rest of the recording. Contrast-agent was automatically injected by the hybrid operating room system with a typical volume of 15 to 20 mL at a rate of 20 to 25 mL per second. The amount of contrast and length of the recording was determined by the surgeon in charge. A purposely induced apnoea was used to minimalize visceral and stent-graft motion during the recording of all series.

For the automatic localization of the renal artery ostia, a total of 95 cDSAs were extracted from the database. Patients with a single visible renal artery on the cDSA were excluded from this study. The dimensions of the cDSAs used for this study had isotropic and anisotropic dimensions ranging from 790 x 1024 to 1904 x 1904 pixels with isotropic pixel sizes ranging from 0.154 x 0.154 $mm^2$ to 0.370 x 0.370 $mm^2$. The cDSA series contained 16 to 31 images per series.

### *3.2.3 Image transformations*

While the cDSA series can be considered as a 2D + time array consisting of 2D images with the time component as the third dimension, we propose to use a 2D approach for the automatic localization of the origins of the renal arteries. This implies that the 2D + time cDSA array is reduced to a single 2D projection, which is then used as the input for the convolutional neural network. The rationale for approaching the renal artery ostia as a 2D localization problem consists of multiple arguments.

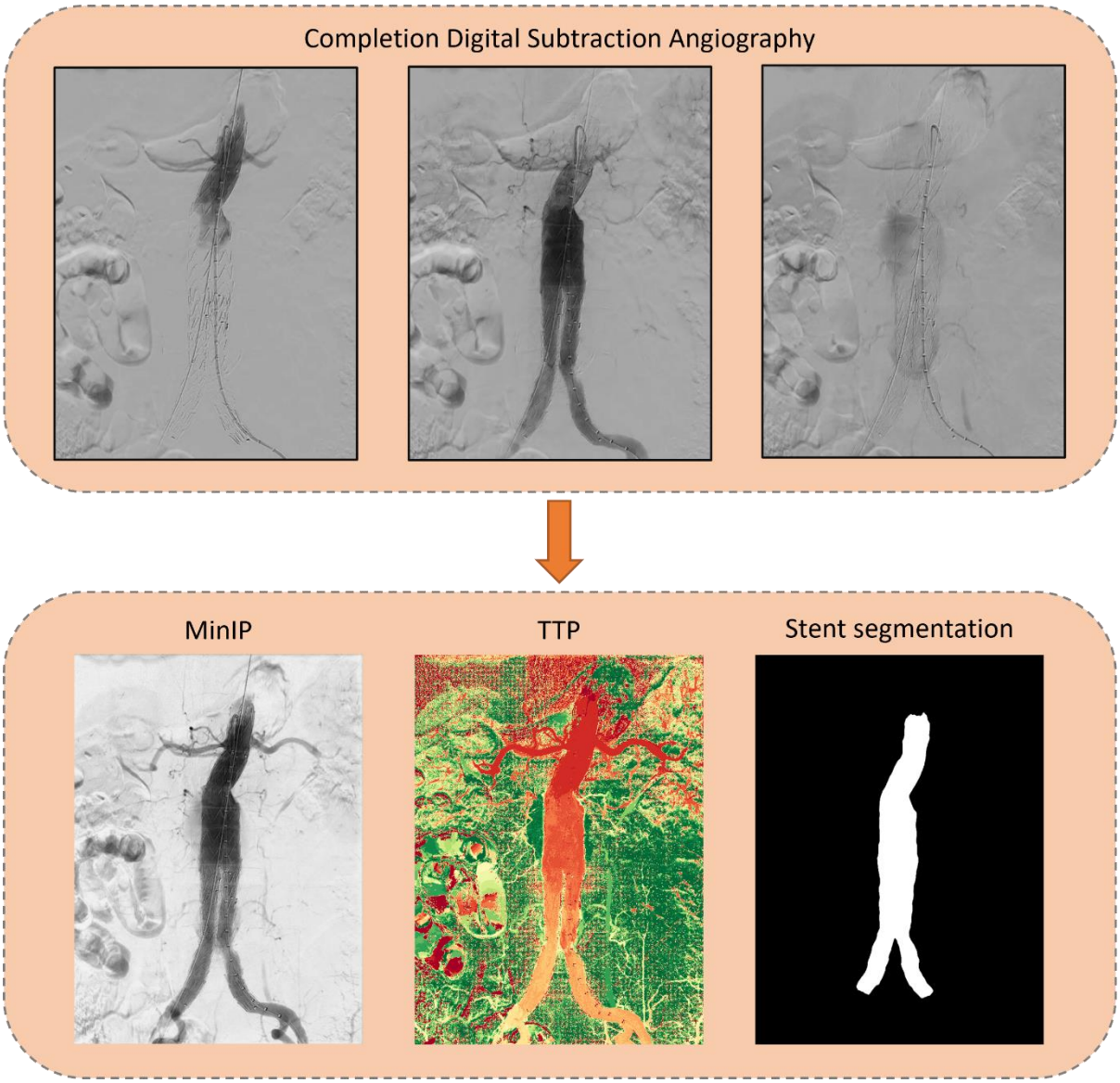


Fig. 6. Top: Images from the completion digital subtraction angiography. Bottom left: 2D Minimum intensity projection (MinIP) illustrating the complete vessel tree. Bottom middle: 2D Time to Peak (TTP) projection illustrating the time from the beginning of the cDSA to the highest density of contrast-agent during the cDSA. Red indicates an early high density of contrast-agent and green indicates a late high density of contrast-agent. Bottom right: 2D stent segmentation predicted by the in-house developed automatic segmentation method (Appendix A).

First, there is an expected inconsistency in the localization of renal artery ostia in the different single 2D frames of the cDSA. Localization of the renal arteries in a series of 2D frames may lead to different coordinates of the renal artery ostia for different frames, which is undesirable in a network from which we desire a single coordinate per renal artery ostium as the outcome. Transformation of the cDSA series into a 2D projection removes this possible inconsistency as the prediction is made on a single image only.

Secondly, by using 2D projections, the renal artery ostium is always visible in the projection, as opposed to using the complete DSA series. In the complete cDSA, the renal artery ostia are not visible throughout the complete series. By reducing the cDSA series into a single 2D projection, the contrast-agent in the

renal artery ostium is always visible. This means that the network does not have to distinguish between frames that do or do not visualize the renal artery ostia.

Reduction and representation of a cDSA series as a single 2D image was achieved by applying a minimum intensity projection (MinIP) along the temporal axis. This way, the projection captures the presence of the contrast-agent throughout the complete series, visualizing the entire vessel tree in a single frame. As we consider the localization of the renal artery as a 2D localization problem by reducing the cDSA into single 2D projections, we assume that there was no movement of the arteries during recording of the cDSA as a result of the purposely induced apnoea. The MinIP projection serves as the main input of the network and is illustrated in Fig. 6.

The consequence of a single MinIP projection of the cDSA series is the loss of temporal information regarding the flow of contrast-agent in the arteries during recording. As this information may be useful for identification and discrimination between target arteries, a Time To Peak (TTP) projection is created to compensate for this loss of temporal information regarding the flow of contrast-agent through the arteries. Every pixel in this 2D projection represents the time in seconds between the start of the cDSA and the highest density of contrast-agent for that individual pixel during recording of the cDSA. This is illustrated in Fig. 6.

The segmentation mask of the stent-graft on the cDSA is added as a third input image to network. The segmentation mask is obtained from an in-house developed deep learning-based automatic segmentation method, see Appendix A. The rationale behind using the segmentation mask as an input for the network is that the annotation of the stent-graft may provide spatial information for localization of the renal arteries. The renal arteries are always located in the proximity of the proximal end of the stent-graft, depending on an infrarenal or suprarenal fixation.

### *3.2.4 Data annotation*

In accordance with previous studies regarding landmark localization, we propose to automatically localize the origins of the renal arteries based on heatmap regression.[43,44] To this end, data annotation of the cDSAs follows a two-step approach to create the ground truth heatmaps. In the first step, the pixel coordinates of the origins of the renal arteries were annotated. Secondly, heatmaps with the dimensions of the corresponding cDSA are created with a Gaussian function representing the location of the renal artery ostia based on the previously annotated coordinates.

The first step of data annotation was performed in MeVisLab[45], a freely available framework for medical image processing. Digital imaging and communications in medicine (DICOM) images of the cDSAs are loaded into a custom-made frame work for data annotation in the software. The x,y-coordinates of the left and right renal artery are annotated on a single frame in the cDSA series. The coordinates are

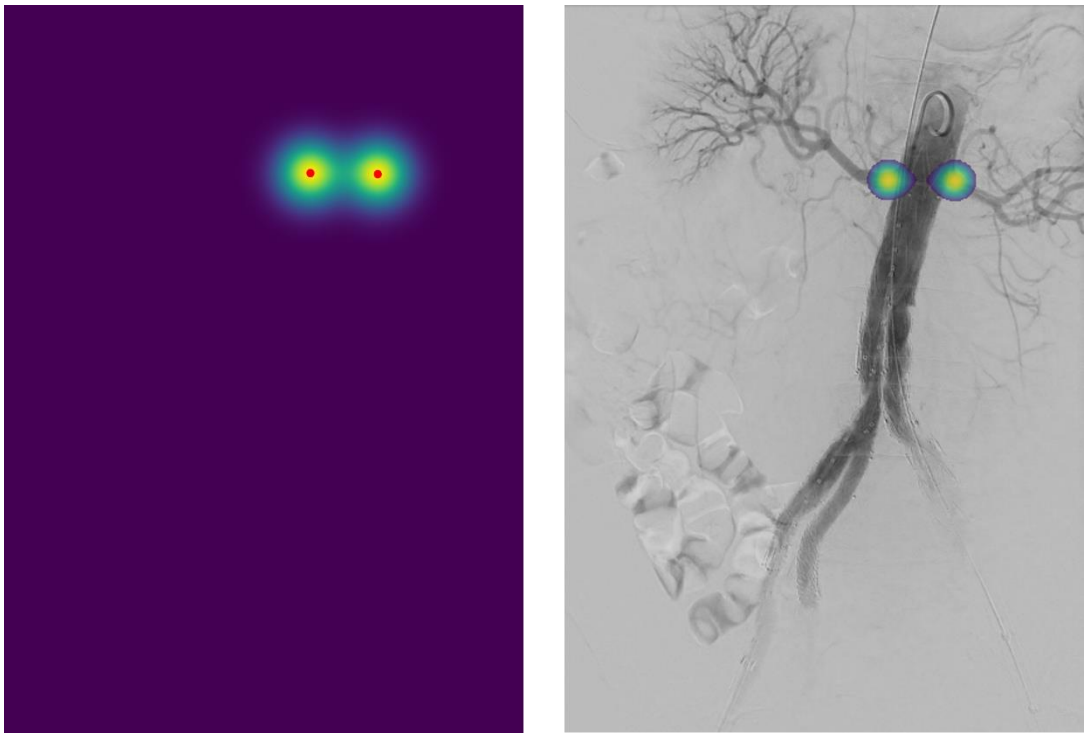

Fig. 7. Left: Ground truth heatmap with Gaussian functions on the location of the annotated ground truth renal artery ostia coordinates (red dots). Right: Gaussian functions projected onto a frame of the cDSA.

converted from world to pixel coordinates and exported as XML-files for further development of the ground truth heatmaps.

In the second step of the data annotation, the pixel coordinates are loaded together with the corresponding DICOM cDSA series of the same patient. The coordinates of the origin of the renal arteries are extracted from the XML-file and split into the right and left renal artery, based on the x- and y-coordinates. An array with the corresponding dimensions of the cDSA is created and the pixel with the locations of the renal artery ostia are marked. A Gaussian filter was applied to obtain a 2D heatmap with a Gaussian function on the location of the renal arteries' origins. The final ground truth heatmap is shown in Fig. 7.

### *3.2.5 Convolutional neural network*

A two-dimensional convolutional neural network (CNN) with a U-Net architecture was trained to automatically predict the location of the origins of the renal arteries from the 2D projections of the cDSA.[42] Whereas a U-Net is commonly used for segmentation, the network is used for regression in this study. The output of the network consists of a single linear output layer, called the prediction heatmap, which illustrates the proximity of a each pixel to the predicted renal artery ostium. As a general pre-processing step, the original cDSA series and ground truth activation maps were resampled to a spacing of 0.4 x 0.4 $mm^2$ using bilinear interpolation. Pixel values were rescaled to values between and 0 and 1 and this was performed in a channel-wise manner in the experiments with multiple input projections.

The database consisted of cDSA series with isotropic and anisotropic proportions and varying dimensions in the x-, and y-direction as they were not recorded in a standardized manner regarding these properties. Following this, the obtained 2D projections are also of different dimensions in the x-, and y-direction. Therefore, the network was trained in a patch-based manner. Image patches with dimensions 512 x 512 pixels were randomly extracted from the 2D projections to train the U-Net. The training images were augmented before patch extraction to create synthetic data to increase the generalizability of the network and partly compensate for the limited number of cDSAs used in this study. The applied data augmentation consisted of horizontal flipping with a probability of 50% and a random rotation around the image center with a maximum deviation of ± 0.25 radians.

A hyperparameter search was performed to find the optimal learning rate and batch size for the network. The final U-Net was optimized for 2500 epochs by minimizing a mean squared error (MSE) loss using an Adam optimizer with a learning rate of 0.0005, dropout rate of 0.2 and batch size of 32. During training, the learning rate was reduced by a factor 10 after 1500 and 2000 epochs. The test images are processed entirely with the original dimensions by the patch-based trained network using a sliding window inference.

The network was implemented using the Pytorch-based framework MONAI[46] and run on a workstation with a NVIDIA GeForce RTX 3090 Graphics Card with 24 GB of memory.

### *3.2.6 Experiments and evaluation of network performance*

The 2D U-Net was trained to predict the location of the renal artery ostia based on a single input consisting of the 2D MinIP projection, a double input comprising the MinIP projection accompanied by the TTP projection, and with a triple input in which the stent segmentation was added to the double input. The effect of the additional input channels on the baseline performance of the single MinIP projection input was evaluated.

For evaluation of the U-Net performance, a nine-fold cross-validation was performed. The database consisting of 95 cDSAs was split into nine folds, each consisting of ten or eleven cDSAs. In nine cycles, eight folds are used as the training set, and the ninth fold is used as a test set to evaluate the performance of the U-Net. During every consecutive cycle, another fold is used as the test set, as illustrated in Fig. 8. The results of the nine cycles are combined to obtain the overall performance of the network.

The performance of the network was evaluated based on the Euclidean distance between the predicted x,y-coordinates of the renal arteries and the ground truth x,y-coordinates. A prediction heatmap was returned by the U-Net illustrating the prediction of the location of the origin of the renal arteries. To obtain the predicted x,y-coordinates of the renal artery ostia, the x,y-coordinates of the two highest activations in the heatmap were selected. A pre-specified minimum distance between the two highest activations was incorporated to avoid two x,y-coordinates that were anatomically unrealistically close to each other. The Euclidean distances between the predicted x,y-coordinates and annotated ground truth

x,y-coordinates of the left and right renal artery were calculated separately and converted into millimetres based on the pixel spacing.

### *3.2.7 Statistics*

The performance and the effect of the number of input channels on the performance of the network was evaluated by comparing Euclidean distance of the localization error and the successful detection rate. The Euclidean distances of the localization error between the prediction and ground truth were reported as the mean with standard deviation or median with interquartile ranges, depending on the normality of the data. The normality of the data was tested using a Shapiro-Wilk test. A one-way ANOVA or Kruskal-Wallis test is used, depending on the normality of the data, to evaluate the difference in localization error between the different input projections. A Student's t-test or Mann-Whitney U test is used to evaluate differences in localization error between the left and renal artery for the same input projections, depending on the normality of the data. Following the example of Noothout et al., we also evaluated the cumulative successful detection rate of the renal arteries based on different thresholds.[47] Statistical analysis was performed using GraphPad Prism version 9.1.0.

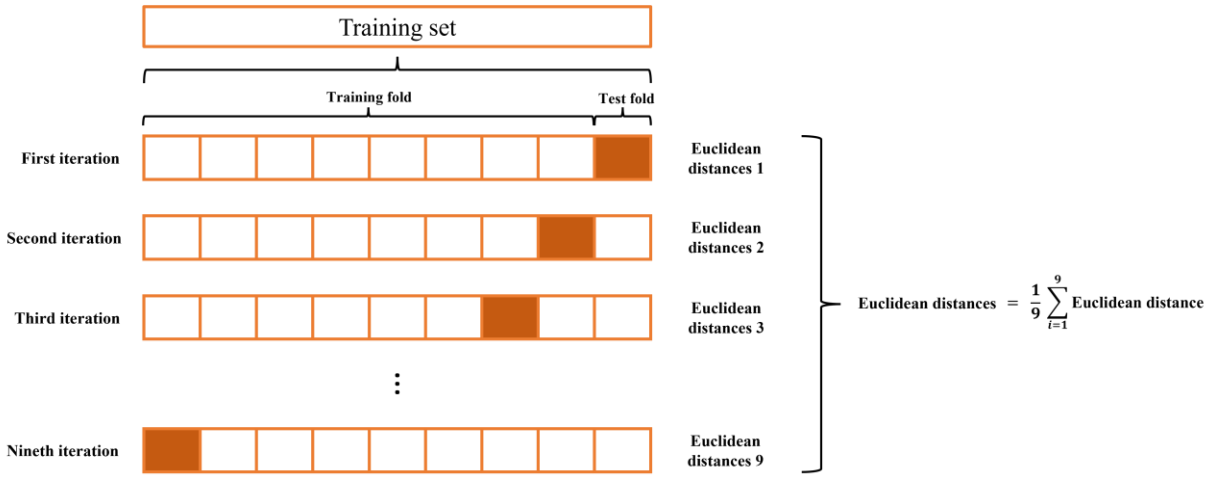


Fig. 8. Illustration of the nine-fold cross-validation for evaluation of the renal artery ostia detection network's performance.

### 3.3 Results

A cross-validation of the U-Net was performed to evaluate the performance of the network for the fully automatic detection of renal artery ostia based on different combinations of 2D feature projections obtained from cDSA's. The predicted x,y-coordinates of the origins of the renal arteries were retrieved from the two maximum activations in the prediction heatmap output of the network. Subsequently, the Euclidean distances between the predicted and ground truth coordinates were calculated in millimetres. Visual results of the prediction of the renal artery ostia by the trained U-Net are shown in Fig. 9 and Fig. 10.

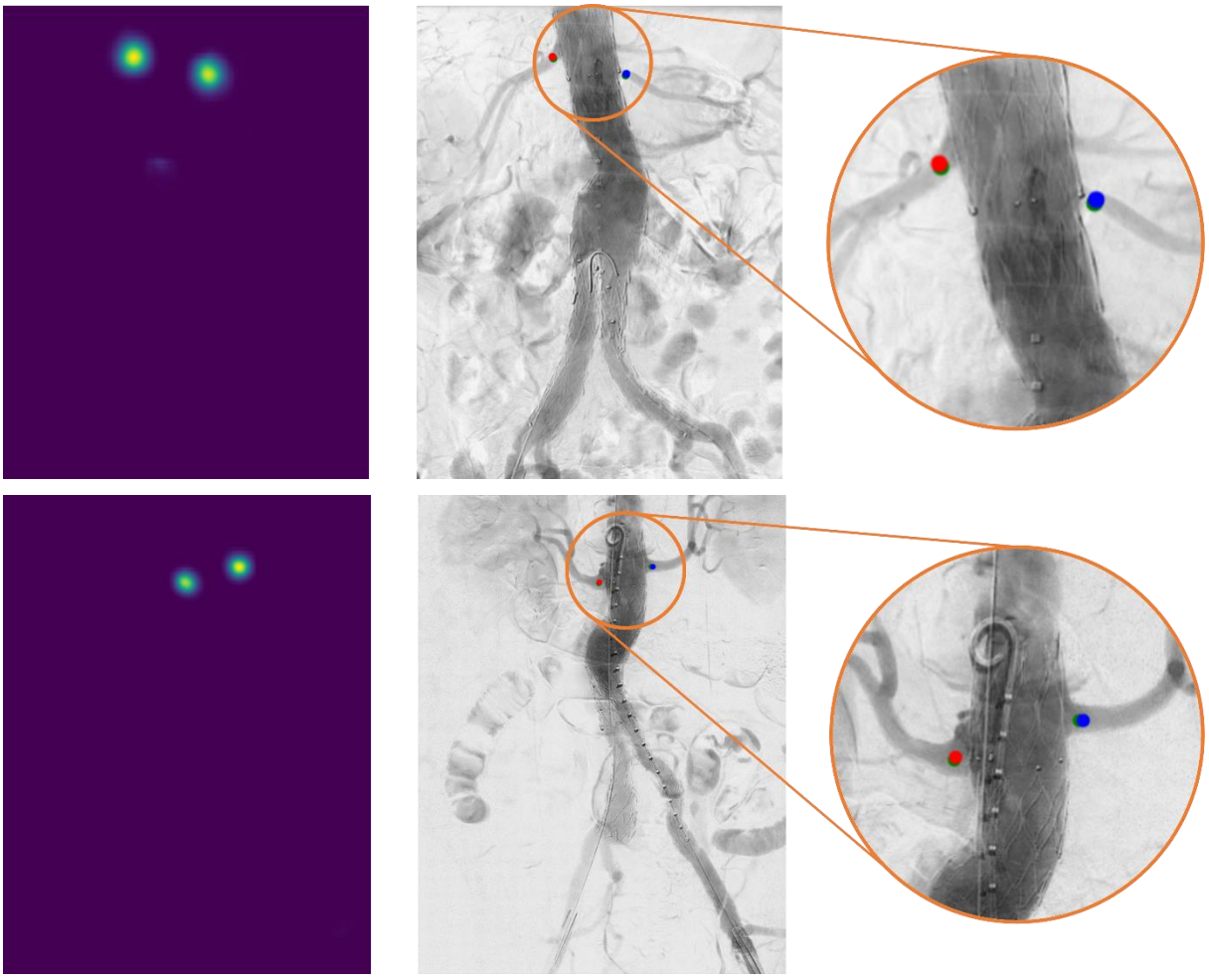

Fig. 9. Qualitative results of the localization prediction of the U-Net with the MinIP projection as a single input. The red dot represents the predicted right renal artery location, the blue dot represents the left renal artery location and the green dots are the ground truth locations of the renal arteries. The heatmap predicted by the U-Net from which the locations were predicted are shown on the left.

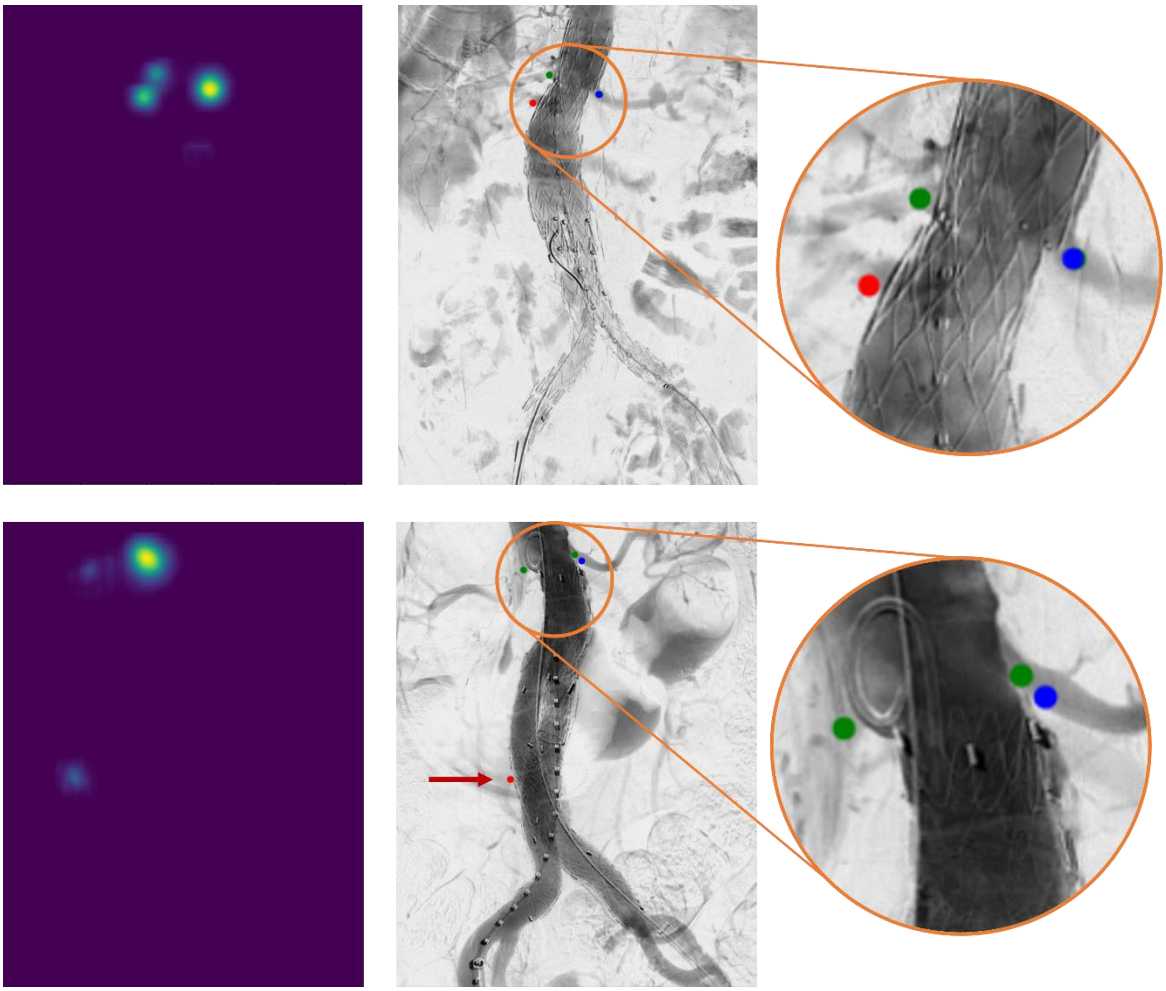

Fig. 10. Qualitative results of suboptimal localization predictions of the U-Net with the MinIP projection as a single input. The red dot represents the predicted right renal artery location, the blue dot represents the left renal artery location and the green dots are the ground truth locations of the renal arteries. The heatmap predicted by the U-Net from which the locations were predicted are shown on the left. Note the major outlier of the right renal artery indicated by the red arrow.

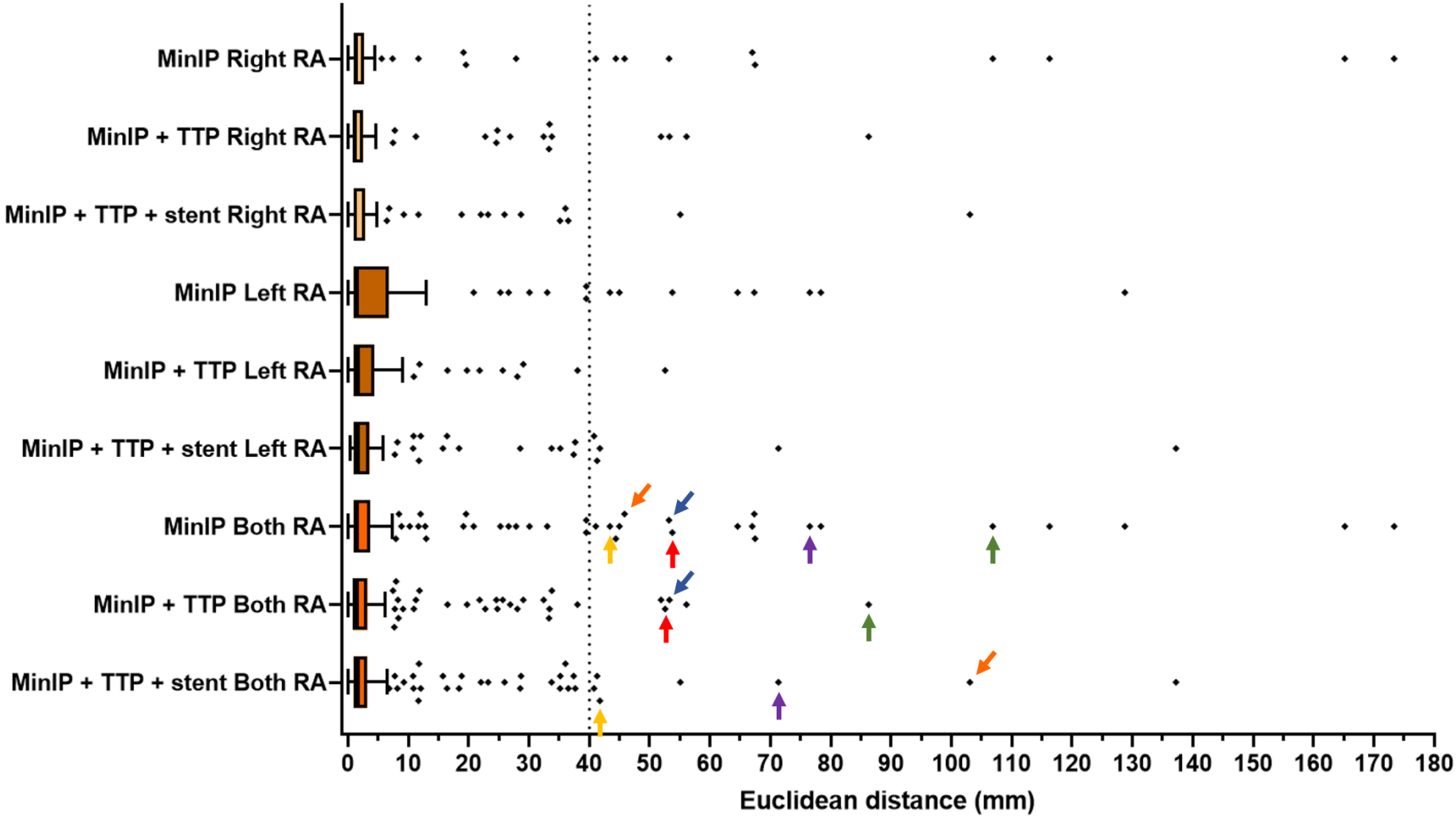


Fig. 11. Boxplots illustrating the distribution of the Euclidean distances between the predicted and ground truth coordinates for the single, double and triple input and divided into right renal artery (RA), left renal artery and both renal arteries. The colored arrows indicate the major outliers on the cDSA of the same patient in different experiments.

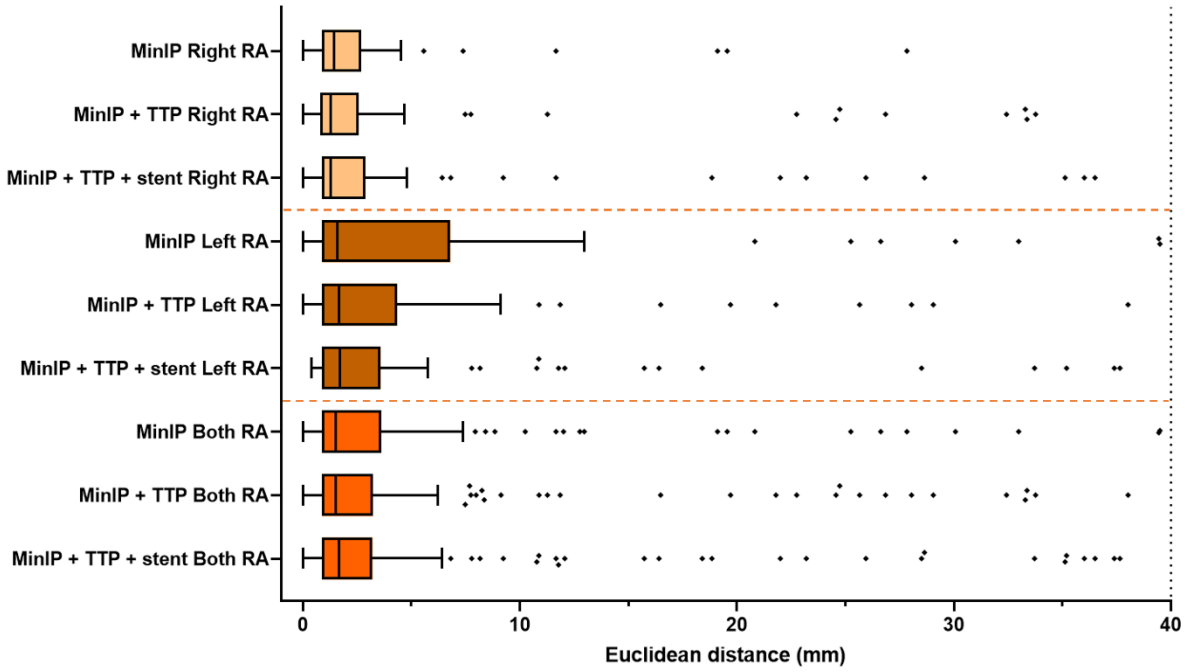


Fig. 12. Zoomed in version of the boxplot in Fig. 11. with a range from 0 to 40 mm.

Figure 11 and 12 show the quantitative results of the cross-validation experiment for the renal artery ostia detection network for the different number of input channels. First, the U-Net with only the MinIP projection as the input image shows the largest outlier (173 mm) and a considerably large number of major outliers (18 outliers above 40 mm). The addition of the TTP projection as a second input decreased the number of major outliers to five. The addition of the stent-graft mask to the MinIP and TTP does not improve the performance of the network in terms of major outliers with 7 predictions with a localization error of more than 40 mm. In total, there were six cDSAs of the same patient with a major outlier in two of the three experiments. There was no cDSA that had a major outlier in all three experiments. Furthermore, there were no cDSAs that had two major outliers for both renal arteries.

The Shapiro-Wilk test on the results of the cross-validation shows a non-normal distribution of the data in the single, dual and triple input group for all renal arteries, right renal arteries, and left renal arteries. For the U-Net with the MinIP as single input, the median Euclidean distance between the x,y-coordinates of the predicted and ground truth location for both renal arteries is 1.52 mm (IQR: 0.89 – 3.605). The median Euclidean distance is 1.44 mm (IQR: 0.89 – 2.68) and 1.60 mm (IQR: 0.89 – 6.790) for the right and left renal artery, respectively. A Mann-Whitney U test showed no significant difference between the results of the left and right renal artery for the single input network ($p=0.360$).

For the U-Net with the MinIP and TTP as input projections, the median Euclidean distance between the predicted and ground truth renal artery ostia locations was 1.52 mm (IQR: 0.87 – 3.22), 1.26 mm (IQR: 0.80 – 2.56) and 1.65 mm (0.89 – 4.33) for both renal arteries, the right renal arteries and left renal arteries, respectively.

| | MinIP | MinIP + TTP | MinIP + TTP + stent | p-value |
|---|---|---|---|---|
| **Right renal arteries** | 1.44 mm (IQR: 0.89 – 2.68) | 1.26 mm (IQR: 0.80 – 2.56) | 1.26 mm (IQR: 0.89 – 2.88) | 0.854 |
| **Left renal arteries** | 1.60 mm (IQR: 0.89 – 6.790) | 1.65 mm (0.89 – 4.33) | 1.70 mm (IQR: 0.89 – 3.58) | 0.958 |
| **Both renal arteries** | 1.52 mm (IQR: 0.89 – 3.605) | 1.52 mm (IQR: 0.87 – 3.22) | 1.65 mm (IQR: 0.89 – 3.21) | 0.853 |

Table 2: Median Euclidean distance between the predicted and ground truth locations of the renal artery ostia. Statistically significant if $p<0.05$

There was no difference between the localization error of the predicted locations between the right renal arteries and the left renal arteries according to the Mann-Whitney U test ($p=0.174$).

The U-Net with three input channels shows a median Euclidean distance between the prediction and ground truth of 1.65 mm (IQR: 0.89 – 3.21) for both renal arteries. The median Euclidean distances for the right renal artery ostia and left renal artery ostia are 1.26 mm (IQR: 0.89 – 2.88) and 1.70 mm (IQR: 0.89 – 3.58), respectively. There is no significant difference between the performance of the U-Net for with regards to the prediction of the right and left renal artery ostia, according to the Mann-Whitney U test ($p=0.260$). The performance results of the U-Net is summarized in Table 2.

Evaluation of the results in Table 2 shows that the performance of the network does not statistically significantly improve in terms of the median localization error if a second or third input channel with the TTP projection and stent segmentation is added to the network according to the Kruskal Wallis test. However, the mean Euclidean distance of the U-Net decreases from 10.8 mm ± 26.41 mm to 5.58 mm ± 11.78 if the MinIP projection is accompanied by the TTP projection and to 6.75 mm ± 16.12 if both the TTP and stent segmentation mask are added as extra input channels. The decreased mean Euclidean distance with multiple input channels is also reflected in the reduced number of major outliers above 40 mm for multiple input channels versus a single input.

The results of the U-Net with a single input and multiple inputs were also evaluated based on the cumulative successful detection rate, shown in Fig. 13. Taking into account the average diameter of the renal artery of 6 mm, and thus a clinically relevant maximum deviation of 3 mm from the central lumen line, the U-Net was capable of successfully detecting 74% of the renal arteries with a single input projection, 72% of the renal artery ostia with two input projections and 75% of the renal artery ostia for the triple input.

Overall, the results show that although there is no significant difference in terms of localization error for the renal artery ostia between the single and multiple inputs to the network, the outcomes of the experiments do suggest an added value of the TTP projection and stent segmentation in terms of the number and magnitude of outliers.

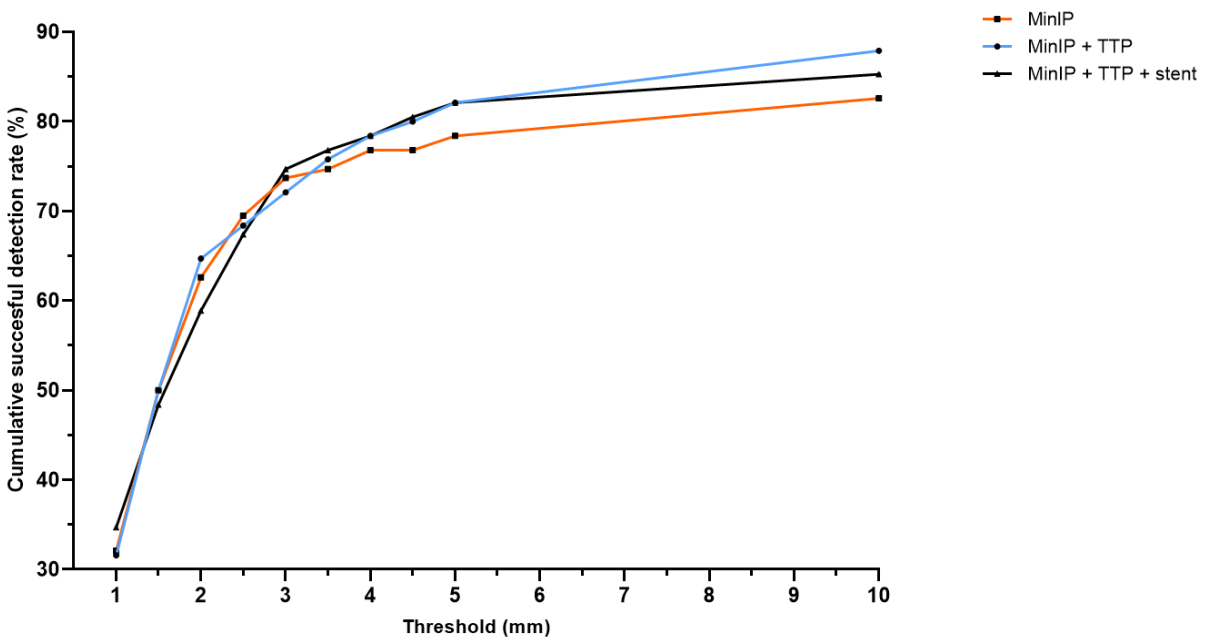


Fig. 13. Cumulative successful detection rate of the renal artery ostia on the cDSA for different thresholds and input projections.

## 3.4 Discussion

In this study, we have demonstrated the possibility to train a convolutional neural network for the automatic localization of renal arteries on a cDSA, using only a 2D MinIP projection and with the addition of a TTP projection and stent segmentation mask. Automatic detection of renal artery ostia in a cDSA series may seem trivial on its own, but in combination with other AI in-house developed artificial intelligence-based models, it can be of added value for automatic and objective analysis of cDSAs. Automatic intraoperative analysis of cDSAs may consist of several aspects, such as evaluation of proximal position of the stent-graft after deployment, detection, and evaluation of endoleaks and evaluation of blood flow dynamics.

### *3.4.1 Interpretation of the results*

Analysis of the results shows us that approximately 75% of the renal arteries can be detected within clinically relevant range of deviation, considering the diameter of the renal artery. A reduced amount of contrast-agent present in the renal artery, together with the reduced clarity of the renal artery contours, seems to negatively influence the performance of the network, as demonstrated by the results in Fig. 10. In the prediction heatmap of MinIP projections with renal arteries that have sufficient contrast-agent and clear contours, we observe activation solely at the site of the renal arteries. In the suboptimal results, the heatmap shows us activations at the site of renal arteries as well. However, activations are higher at the site of contrast-agent induced artefacts in the intestines. The amount of contrast-agent, together with the clarity of the renal artery contours seem to be an important factor for the performance of the network. The clarity of the renal arteries and the amount of contrast-induced artefacts may be the result of

movement during acquisition of the cDSA, despite the induced apnea. Furthermore, a lateral course of the superior mesenteric artery in close proximity of the renal arteries led to incorrect localizations in some cases.

The temporal information regarding blood flow and appearance of the visceral arteries throughout the cDSA series is lost in a 2D MinIP projection. The renal arteries are visualized at the beginning, and thus before the majority of other arteries, of the cDSA recording as the pigtail catheter for contrast-agent injection is placed proximal to these arteries. The time component in the cDSA series may be important to discriminate between other arteries and contrast-induced artefacts in the intestines. To partly regain this potentially important temporal information, a TTP projection was added as a second input channel to compensate for the loss and possibly increase the performance of the network. However, despite less major outliers and a reduced average Euclidean distance between the predicted and ground truth coordinates, there no significant advantage nor disadvantage of adding this second input channel.

The addition of the stent-graft segmentation as a third input next to the MinIP and TTP projections was expected to further improve the accuracy of the network. The stent-graft segmentation provides the network with indirect spatial information on where the renal arteries can be expected in the frame of the cDSA. For the EVAR procedures, the renal arteries are always located proximally or distally to the proximal end of the stent-graft, depending infrarenal or suprarenal fixation. The median and mean localization error did not significantly change after addition of the stent-graft segmentation. The accuracy of the stent-graft segmentation mask might influence the added value that the stent segmentation mask input has on the performance of the network. However, evaluation of the stent-graft masks for the major outliers shows a good visual results of the stent-graft segmentation masks that were used as input to the network.

The purposely induced apnea minimizes the movement due to ventilation of the patient but the cardiac cycle may still induce minor movements of the arteries. However, during this experiment, with the use of 2D projections from a 3D array, the assumption was made that there was no movement of the visceral arteries during the recording of the cDSA. Even though minor movements may be possible, we expect minor influence hereof in the results of the experiments. Another note that needs to be made is the use of millimeters to express the Euclidean distance of the localization error. The distance between the X-ray source, patient and detector influences the scale of the patients anatomy on screen. Therefore, the Euclidean distances can be used as an indication of the localization error but cannot be considered as the true error.

### *3.4.2 Comparison with literature*

Automatic localization of landmarks with deep learning has been reported extensively in previous studies.[43,44,47] To the best of our knowledge, this is the first study that uses deep learning for automatic landmark localization on cDSA series during EVAR. Payer et al. investigated several CNN architectures

for automatic localization of multiple landmarks in 2D X-ray radiographs of the hand. Their approach is closest to our approach regarding the network architecture, creation of ground truth heatmaps, and experimental setup. Using a U-Net, they achieved a median localization error of 0.68 mm on 37 landmarks in the radiographs, which is slightly less than half of our median localization error.[43] In contrast to our data set of 95 images, they were able to optimize their network based on a data set with close to 900 images, while their 2D X-ray images are also expected to have less variation in terms of anatomy, background structures and artefacts. Noothout et al. proposed a global-to-local localization approach in which regression and classification were performed simultaneously to localize landmarks in Cephalometric X-rays. With a more sophisticated approach compared to ours, they report a median localization error of around 1 mm with a data set consisting of 400 radiographs.[47]

### *3.4.3 Limitations of the study*

Several limitations exist within this study. Deep learning networks are capable of objectively processing and learning from large amounts of data that may exceed the exposure of a single human. The current study was performed based on a limited number of cDSAs from a single center. Expansion of the data set is expected to substantially increase the performance of the model and reduce the median Euclidean deviation and increase the successful detection rate at lower thresholds.

In this study, there was a variability in the image dimensions of the cDSA's. We used a patch-based approach during training of the U-Net as the network requires similar input dimensions for all images. Resizing of the images to the same dimensions induces distortion of the image proportions and was therefore not considered a viable option. A limitation of using a patch-based approach is the loss of spatial information with regard to the position of the renal arteries on the image, which may be beneficial for localization of the renal artery ostia. Furthermore, since the renal arteries are mostly located at the top of the cDSA images, random patch extraction from the 2D projections during training may induce and imbalance in the number of patches with and without renal arteries. This is because pixels towards the boundaries of the images are extracted less during patch extracting than pixels towards the center of the image. A weighted patch extraction was also considered, however, this would lead to a bias in the network as it may expect a renal artery ostium in every patch which is not favorable during patch-based sliding window inference evaluation of a 2D projection with original image dimensions.

A last limitation of this study is the lack of external validation. The application of cross-validation in our study validated the performance of the network internally. However, external validation on an independent data set needs to be performed to test the performance of the network in different institutions. The current study only used cDSA images acquired in a Philips hybrid operating room, but the performance on cDSAs acquired at other institutions using a system of a different manufacturer is unknown.

### *3.4.4 Future perspectives*

For future work on this study, there are several aspects that could be focused on to further improve the accuracy of the network and increase the successful detection rate at lower thresholds. Cina et al., proposed a deep learning method for automatic detection of vertebral corners on 2D sagittal X-ray images, a study that in terms of imaging modality relates to this paper.[44] In contrast to our method, they use a two-step approach; the first CNN localizes the vertebra and identifies the vertebral corners, in the second step, the localization is refined in a smaller patch of the image. The application of this two-step approach may be useful in our study, as the highest activation is always in close proximity to the renal arteries. Using a cropped images based on the location of the highest activation, the refined localization of the origin of the renal arteries is less influenced by other arteries and contrast-induced artefacts.

## 3.5 Conclusion

In this study, we have demonstrated the feasibility of using a U-Net for the automatic localization of the renal arteries on a completion digital subtraction angiography during EVAR. Experiments have shown that the addition of a second input channel to compensate for the lost temporal information in the 2D MinIP projection leads to fewer major outliers but does not significantly improve the localization error and successful detection rate of the network. The addition of the stent-graft segmentation did not improve the localization error or reduce the number of major outliers.

Chapter 4

# Deep learning-based identification of lumbar arteries on completion digital subtraction angiographies during Endovascular Aneurysm Repair

## 4.1 Introduction

Since its introduction, Endovascular Aneurysm Repair (EVAR) has increasingly gained popularity and established itself as the first choice of treatment for elective repair of an infrarenal abdominal aortic aneurysm (AAA) as opposed to invasive open surgical repair.[41] The minimally invasive approach of EVAR compared to open surgical repair has led to short-term benefits in terms of peri-operative survival, severe complications and hospital stay.[8,11] Despite these early advantages of EVAR, the endovascular approach is being plagued by a significantly higher reintervention rate due to complications, of which endoleaks make up the majority.[12]

Endoleaks make up the majority of indications for a secondary intervention after EVAR, with type 2 endoleaks being the most encountered type.[12,48] Type 2 endoleaks are defined as a retrograde flow of blood into the peri-graft volume of the aneurysm sac after deployment of the stent-graft. Sources of the retrograde flow are the inferior mesenteric artery, accessory renal arteries, intercostal arteries and lumbar arteries.[49] The number of patent lumbar arteries has been demonstrated as a risk factor for the development of a type 2 endoleak.[38] Furthermore, Seike et al. has also demonstrated that the number of patent lumbar arteries is significantly correlated with a higher risk of the development of persistent type 2 endoleaks and sac enlargement after EVAR.[35]

The first encounter with a possible endoleak is during the completion angiogram at the end of an EVAR procedure; a digital subtraction angiographic technique that visualizes the blood flow with the injection of a contrast-agent. Clinical decision-making is based on the completion angiogram. However, the evaluation of the completion angiogram is performed by subjective visual inspection and therefore user-dependent. Detection and classification of endoleaks on the completion angiogram is not always trivial and shows moderate inter-observer agreement between medical specialists as we have seen in an inter-observer agreement study by Smorenburg et al. In this study, medical specialists independently scored completion angiograms on the presence and type of endoleaks. As the evaluation of the completion angiogram remains subjective with moderate inter-observer agreement, automatic objective intraoperative image analysis such as automatic detection and classification of endoleaks can potentially increase the inter-observer agreement between medical specialists and improve the outcome of the procedure.

Automatic identification of lumbar arteries can help in the objective intraoperative evaluation of completion angiograms. Although acquisition of the cDSA is performed in a single projection plane, the study in chapter 3 has shown that over-projection of the lumbar arteries by the contrast-agent in the aortic lumen is very rare. Presence of lumbar arteries in the proximity of detected endoleaks can aid in the classification of the endoleak. Therefore, the aim of this study is to use a deep learning-based approach for the automatic identification and localization of lumbar arteries on a completion angiogram and investigate the influence of additional input channels on the performance of the network. We hypothesize

that the deep learning network is capable of identifying lumbar arteries on the cDSA and that additional input channels increase the performance of the network.

## 4.2 Methods

### *4.2.1 Proposed pipeline*

This study aims to automatically identify the presence of lumbar arteries on a completion digital subtraction angiography acquired at the end of an EVAR procedure. A convolutional neural network with a U-Net[42] architecture was trained to return a prediction heatmap that shows the location and course of the lumbar arteries. Two-dimensional (2D) minimum intensity projections (MinIP) obtained from cDSA series were used as the main input to the network and extra 2D feature projections extracted from the cDSA series were added to the main input to investigate the influence hereof on the performance of the network. The network performance for the single and multiple input channels is evaluated based on precision-recall curves.

### *4.2.2 Data*

To train a network for identification of lumbar arteries on a cDSA, we used a data set consisting of 70 completion digital subtraction angiographies (cDSA) of patients that were treated for an infrarenal abdominal aortic aneurysm using EVAR. All cDSAs were acquired using a C-arm X-ray system in a Philips Azurion FlexMove 7 C20 Hybrid Operating Room (Philips Healthcare, Best, The Netherlands) in the Amsterdam University Medical Centers, location VUmc, between April 2017 and April 2020. Images were retrospectively extracted from a prospectively maintained database with approval of the local medical ethical review committee.

A multiphase acquisition protocol was used for the acquisition of the cDSAs: during the first phase, images are acquired at a frame rate of 3 frames per second (fps) for 6 seconds, during the second phase, images are acquired at 1 fps for 5 seconds and images are acquired at 0.5 fps during the third phase for the rest of the recording. Contrast-agent was automatically injected by the hybrid operating room system with a typical volume of 15 to 20 mL at a rate of 20 to 25 mL per second. The amount of contrast and length of the recording was determined by the surgeon in charge. A purposely induced apnoea was used to minimalize visceral and stent-graft motion during recording of all series. The dimensions of the cDSA and isotropic pixel sizes were 790 x 1024 to 1904 x 1904 pixels and 0.154 x 0.154 $mm^2$ to 0.370 x 0.370 $mm^2$, respectively. The number of frames in the cDSA series ranged from 16 to 43.

### *4.2.3 Image transformations*

In this study, we propose to face the identification of lumbar arteries on a cDSA with a 2D approach, even though the cDSA series is a 2D + time array consisting of multiple consecutively acquired 2D images which can be seen as a 3D array. Therefore, the 2D + time array consisting of multiple frames of the cDSA is reduced to a single 2D frame which is used as input for the convolutional neural network. The

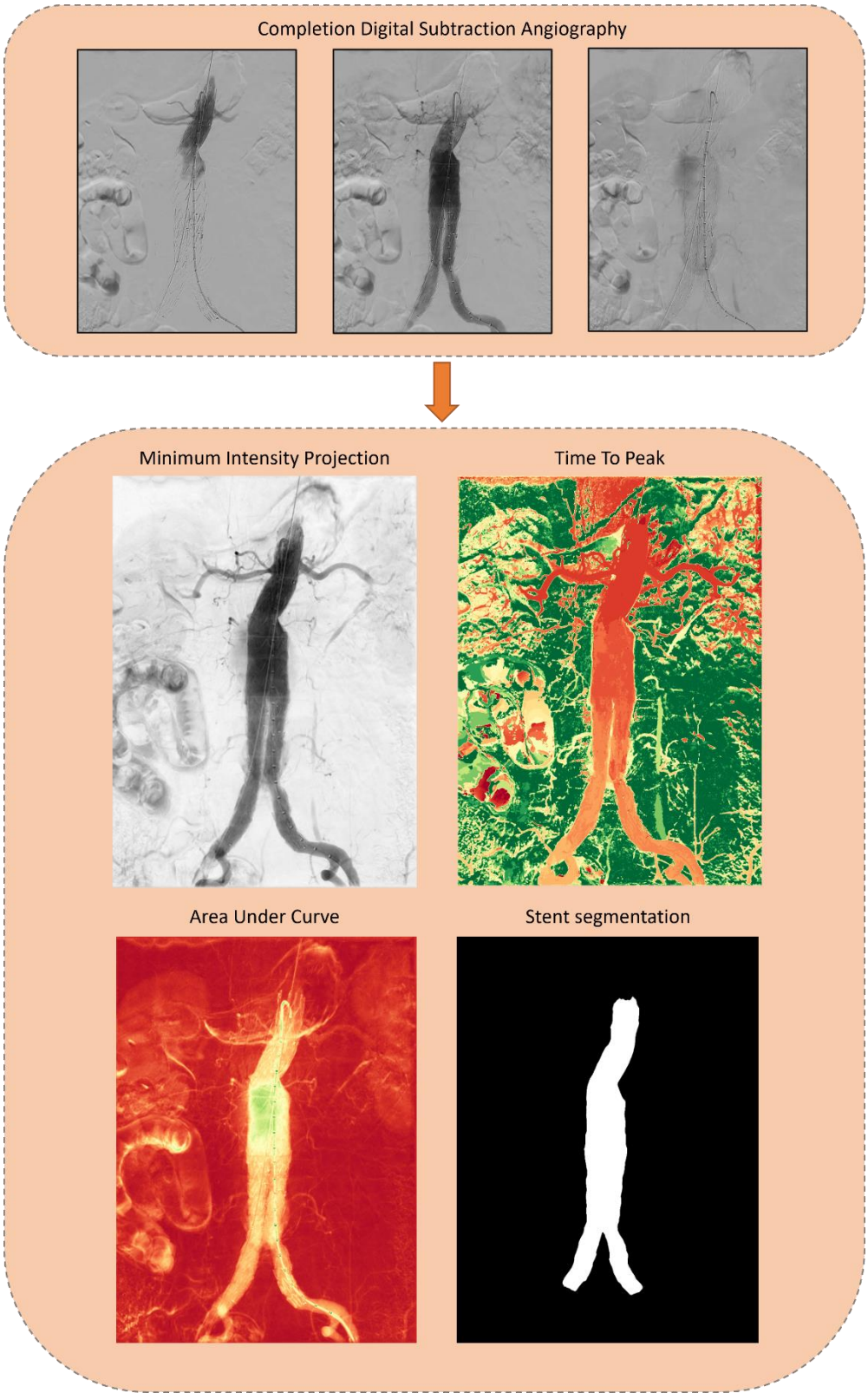


Fig. 14. Illustration of the 2D feature projections obtained from the completion digital subtraction angiography. For the Time To Peak, red indicates an early high density of contrast-agent and green indicates a late high density of contrast-agent. In the Area Under Curve projection, red indicates a low area (less contrast-agent) and green indicates a large area under the curve (more contrast-agent) in each pixel during the cDSA. The stent segmentation is obtained from the deep learning stent-graft segmentation model (Appendix A).

different 2D input projections are shown in Fig. 14. Several arguments can be given for using 2D projections as input for the neural network as opposed to the complete 3D array of the cDSA.

First, a potential advantage of using 2D projections as opposed to the entire 2D + time cDSA series as input for the network is that the network does not have to distinguish between images on the visibility of the lumbar arteries. Second, transformation of the cDSA series into a single projection visualizes the entire lumbar artery that would be visible over the course of the cDSA series. In individual frames of the cDSA series, only a small part of the lumbar artery may be visible as the rest of the artery has not been filled with contrast-agent. This is achieved by application of a Minimum intensity projection (MinIP) along the temporal axis of the cDSA series. The MinIP visualizes the highest density of contrast-agent reached for each pixel during the cDSA and therefore visualizes the entire vessel tree that was filled with contrast-agent. Identification of lumbar arteries is expected to be more accurate in case of maximum visibility of the lumbar artery which is achieved by this single 2D projection.

A Time to Peak (TTP) projection illustrates the time between the start of the cDSA acquisition and the time at which the contrast-agent density in the artery reaches its highest point. This 2D projection is therefore a representation of how fast a given artery is visible after injection of the contrast-agent. As the lumbar arteries at the height of the stent-graft are mostly visible in a later stage of the cDSA series due to the inflow of contrast-agent through collateral arteries, the information captured in this projection may be of added value to discriminate between multiple visible arteries. The temporal information on the visibility of the arteries in the cDSA series is lost by using the MinIP as the main input, and this TTP projection compensates for this loss of flow dynamic information in the cDSA.

The total amount of contrast-agent within a given artery during the acquisition of the cDSA can be captured in an Area Under the Curve (AUC) projection. After stent-graft deployment, the contrast-agent reaches the lumbar arteries at the height of the stent-graft through collateral arteries, while other arteries such as the renal arteries and superior mesenteric artery including branches have a more direct inflow of contrast-agent. The contrast-agent concentration in the lumbar arteries may therefore be lower due to dilution and thus a lower total amount of contrast-agent flows through the lumbar arteries. The AUC projection may help in correctly identifying the lumbar arteries from other arteries based on the total amount of contrast-agent that within the artery during acquisition of the cDSA.

Lumbar arteries are commonly present alongside the stent-graft. The position of the stent-graft on the cDSA can provide spatial information to the network on the expected locations of the lumbar arteries. For this study, a stent-graft segmentation mask is obtained from all cDSA series using a deep learning network for automatic stent-graft segmentation on cDSA images (Appendix A).

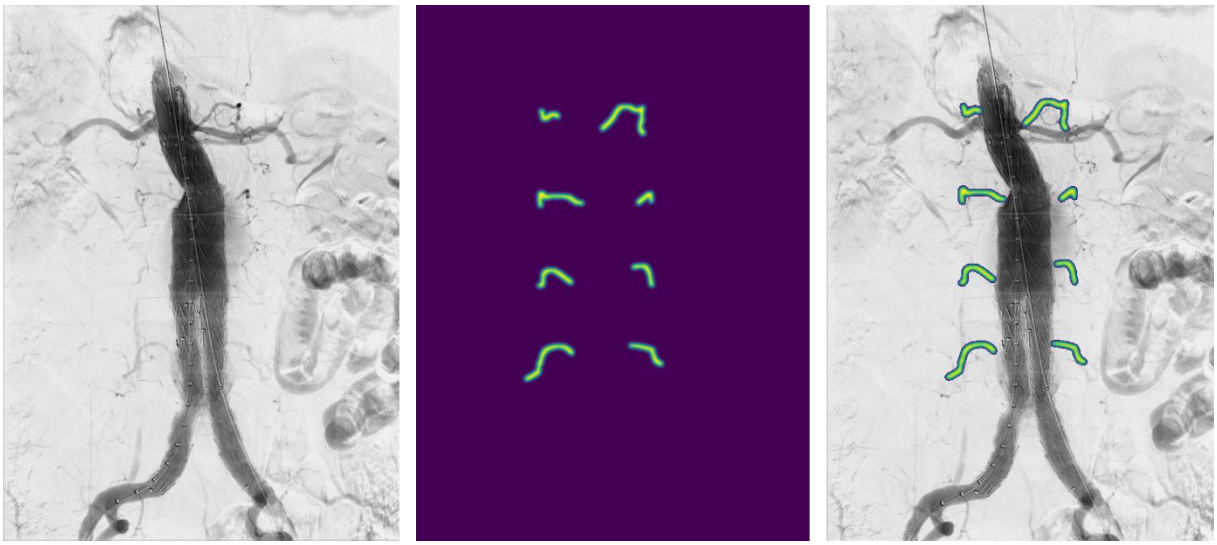

Fig. 15. Left: MinIP projection of the cDSA. Middle: Ground truth heatmap with Gaussian filtered lines that represent the location and course of the lumbar arteries. Right: Ground truth heatmap projected onto the MinIP projection.

### *4.2.4 Data annotation*

Annotation of the ground truth heatmaps for automatic identification of the lumbar arteries is partly based on the data annotation methods for landmark localization problems using heatmap regression as illustrated by earlier studies and the annotation of the renal artery ostia in chapter 4.[43,44] A similar two step approach as for the data annotation of the renal artery ostia is used for the lumbar arteries. Lumbar artery ground truth heatmaps differ in the fact that a Gaussian filtered line is used to represent the visible course of the artery on the cDSA as opposed to a Gaussian function representing a single landmark at the ostia.

For the first step, a custom framework was setup in MeVisLab[45] for annotation of the lumbar arteries. The lumbar arteries were identified on the cDSA with the guidance of the preoperative CTA and maximum intensity projection of the preoperative CTA. For this, a similar setup was used as described in chapter 3. A central lumen line was drawn over the course of the visible lumbar artery on a single frame in the cDSA. The coordinates of the central lumen line were subsequently converted from world to pixel coordinates. In the next step, the central lumen lines of the lumbar arteries were dilated and the entire ground truth frames were processed with a Gaussian filter. A maximum intensity projection was performed along the temporal axis of the ground truth frames to obtain a 2D ground truth heatmap with all labelled lumbar arteries as illustrated in Fig. 15.

### *4.2.5 Convolutional Neural Network*

For the automatic identification of lumbar arteries on the cDSA based on the 2D projections obtained from the cDSA, a two-dimensional convolutional neural network with a U-Net architecture was trained.[42] The output of the U-Net is a single linear output layer that represents a prediction heatmap illustrating the

distance of each pixel to the central lumen line of a lumbar artery, as opposed to the more known binary output for segmentation tasks.

The U-Net was trained in a patch-wise manner with image patches of 512 x 512 pixels randomly extracted from the 2D projections. Data augmentation was applied prior to image extraction to create additional synthetic data to increase the generalizability of the network and compensate for the limited number of images available for the study. Image patches were randomly flipped horizontally with a probability of 50%. As a general pre-processing step prior to input to the network, the 2D projections were resampled to pixel sizes of 0.2 x 0.2 $mm^2$ using bilinear interpolation and rescaled to a pixel range of [0, 1] and in a channel-wise manner in case of multiple input projections.

The network was optimized using a mean squared error (MSE) loss together with an Adam optimizer algorithm. Hyperparameters of the U-Net were a learning rate of 0.001, dropout rate of 0.2 and a training batch size of 64 image patches. The training was performed for 3000 epochs with a reduction of the learning rate by a factor 10 after 2000 and 2500 epochs. A sliding window inference was used to process the test images with the original dimensions by the patch-based trained network.

The network was implemented using the Pytorch-based framework MONAI[46] and run on a workstation with a NVIDIA GeForce RTX 3090 Graphics Card with 24 GB of memory.

### *4.2.6 Post-processing*

The prediction heatmap outputs of the U-Net are post-processed to obtain a binary map for the prediction of the presence of a lumbar artery on the cDSA. First, the prediction heatmap is post-processed using a grayscale closing which consists of subsequent grayscale dilation and grayscale erosion. This step is performed to connect multiple activation segments of the same lumbar artery in the prediction heatmap. A binary output image is subsequently obtained from the prediction heatmap by thresholding based on pixel intensity. Small components in the binarized image are removed based on a minimum area that the components need to have in order to be considered a possible lumbar artery. Smaller components than the minimum area were considered as artefacts. The remaining components are dilated to expand the predicted region of the lumbar artery presence. The ground truth heatmaps are similarly thresholded to obtain a binary ground truth mask. The workflow of the post-processing steps are illustrated in Fig. 16.

### *4.2.7 Experiments and evaluation of network performance*

The U-Net was trained for identification of lumbar arteries based on different combinations of input projections. In the first experiment, only the MinIP projection was used as the input image for the network; in the consecutive experiments, the MinIP projection was accompanied by the TTP, AUC or stent-graft segmentation mask.

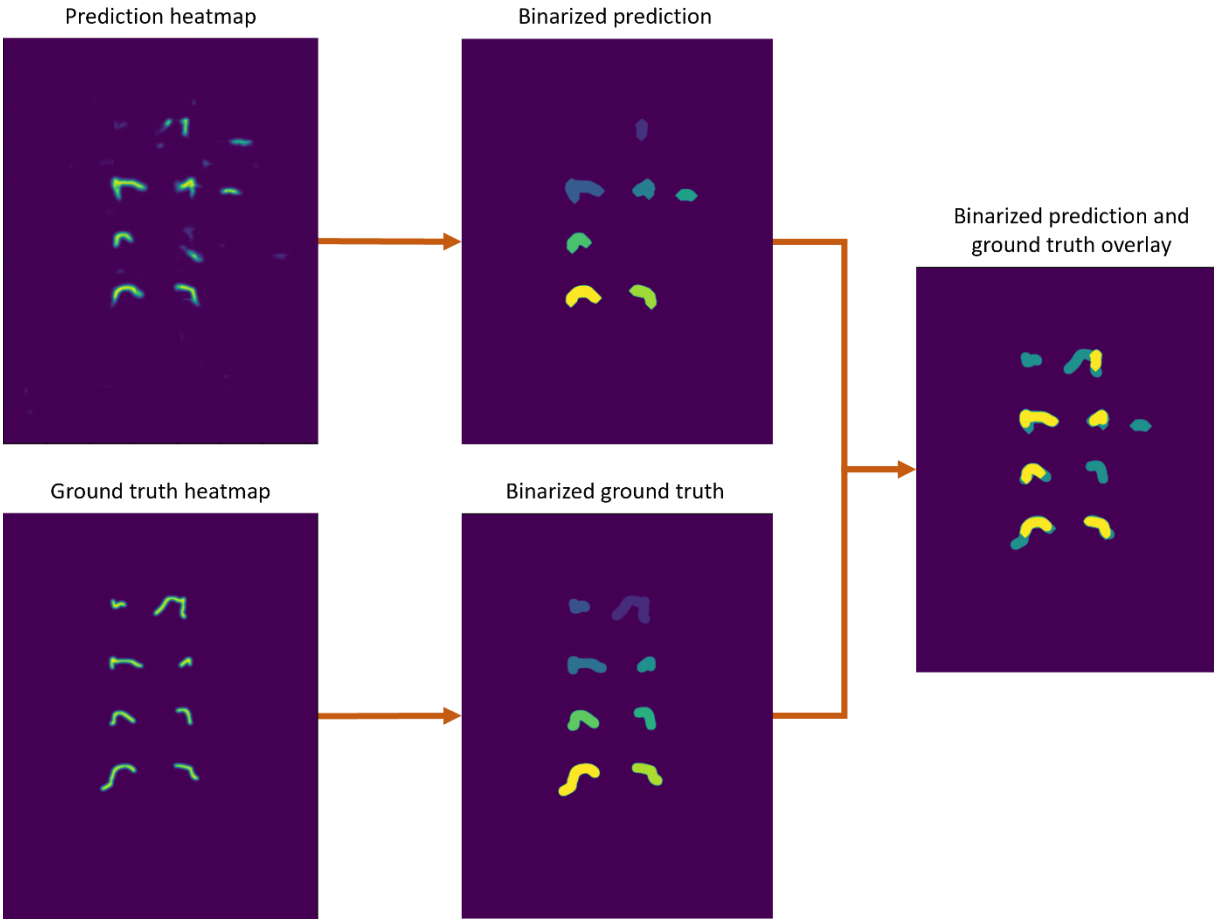


Fig. 16. Flowchart of the post-processing steps. A binarized prediction is obtained from the prediction heatmap by grayscale closing, thresholding based on pixel intensities, selection of components based on the area and dilation of the prediction components. The binary ground truth is obtained by thresholding the ground truth heatmap based on pixel intensity. The overlay of the binarized prediction and ground truth is used for calculation of the TP, FP and FN.

A seven-fold cross-validation of the U-Net was performed for the each of the four experiments. In the cross-validation, the data set was split into seven folds, each consisting of ten cDSAs from which the projections are obtained. In every cycle, six folds are used as the training set, and the seventh fold is used as the test set.

The performance of the network for each combination of input projections is evaluated based on a precision-recall curve. For every fold in each of the cross-validations, the prediction heatmaps are post-processed as described in section 5.2.6 and illustrated in Fig. 16. to obtain the binarized predictions. The binarized prediction is compared with the binarized ground truth and the true positive (TP), false positive (FP), and false negative (FN) predictions are determined. Overlap between components in the binarized prediction and binarized ground truth were considered as a correct prediction, thus TP. A component in the binarized prediction without an overlap of a component in the binarized ground truth was considered as FP and a component in the binarized ground truth without a component at the same location in the binarized prediction was considered as FN.

For every fold in the cross-validations, the TP, FP and FN are determined for ten different thresholds of pixel intensity values. For each experiment, a precision-recall curve is created based on the TP, FP and FN at different thresholds for the complete cross-validation on 70 cDSAs. The area under the curve in the precision-recall curve ($AUC_{PR}$) is determined to obtain the overall performance of the network with the MinIP, MinIP + TTP, MinIP + AUC, and MinIP + stent-graft segmentation mask as input projections to the network. The $AUC_{PR}$ for the different experiments are compared to determine the combination of input projections for the best performance of the network.

### 4.3 Results

A seven-fold cross-validation was performed on the full data set of 70 cDSAs to evaluate the performance of the trained U-Net for the automatic identification of lumbar arteries on the completion angiogram. The performance of the network was evaluated based on the $AUC_{PR}$ of the cross-validation for the different combinations of input projections. Fig. 17 shows the qualitative results of the prediction of the presence and location of lumbar arteries on the cDSA.

The influence of thresholding on the final prediction of the presence and location of lumbar arteries on the cDSA is illustrated in Fig. 18. A lower threshold leads to a higher number of TP predictions and low number of FN, but also leads to a higher number of FP predictions. Increasing the threshold value for creating a binary prediction leads to a lower number of FP predictions but also decreases the number of TP predictions and increases the number of FN. Fig. 19 illustrates another final prediction if a low threshold is used for creating the binary prediction. The FP activations in the prediction heatmap are commonly located on the superior mesenteric artery (SMA) and the SMA branches.

For each of the experiments with a different combination of input projections, a precision-recall curve was created based on the results of the cross-validation. The precision-recall curves are illustrated in Fig. 20. The $AUC_{PR}$ are 0.59, 0.61, 0.67 and 0.62 for the MinIP, MinIP + TTP, MinIP + AUC, and MinIP + stent segmentation mask input projections, respectively. The addition of an extra input channel with a 2D feature projection showed improvement of the $AUC_{PR}$ in all cases. The addition of the AUC projection to the MinIP as a second input projection showed the largest increase of the $AUC_{PR}$.

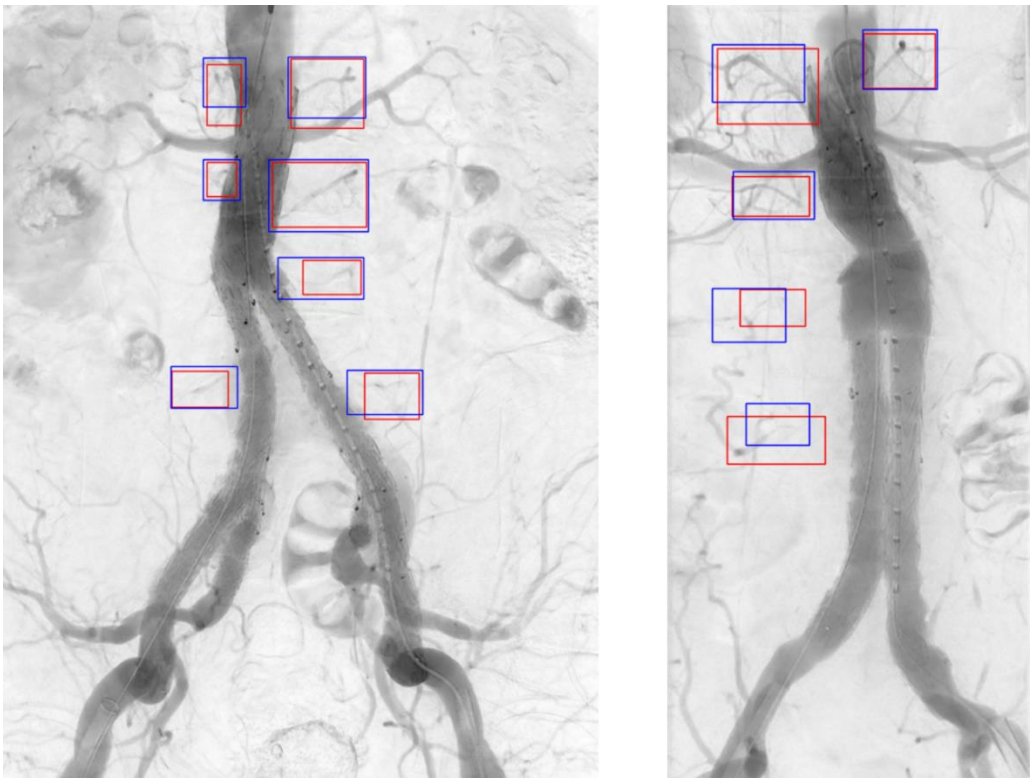

Fig. 17. Qualitative results for the identification of lumbar arteries on the cDSA. The results shown are generated by the U-Net that was trained based on the MinIP and AUC projections. Bounding boxes are created around the binary prediction and binary ground truth. The red boxes represent the ground truth, the blue boxes represent the prediction of the network.

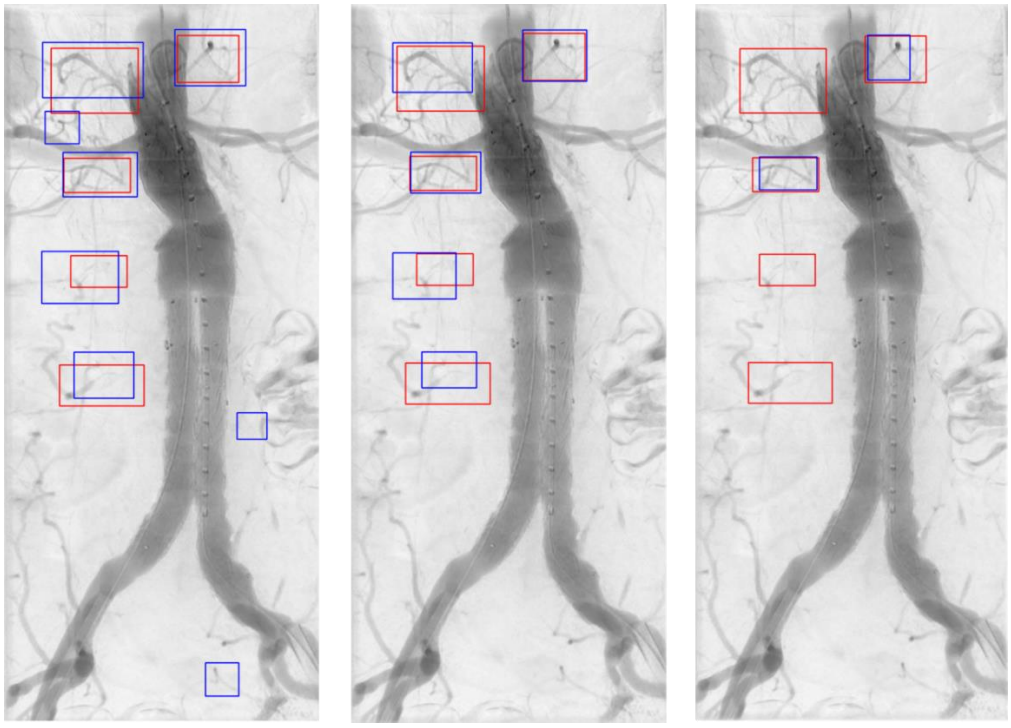

Fig. 18. Illustration of the prediction by the U-Net trained on the MinIP + AUC projections for different thresholds of pixel intensity during post-processing of the prediction heatmaps. A lower threshold (left image) results in a higher rate of FP predictions, while a higher threshold (right image) results in a higher rate of FN predictions. The red boxes represent the ground truth, the blue boxes represent the prediction of the network.

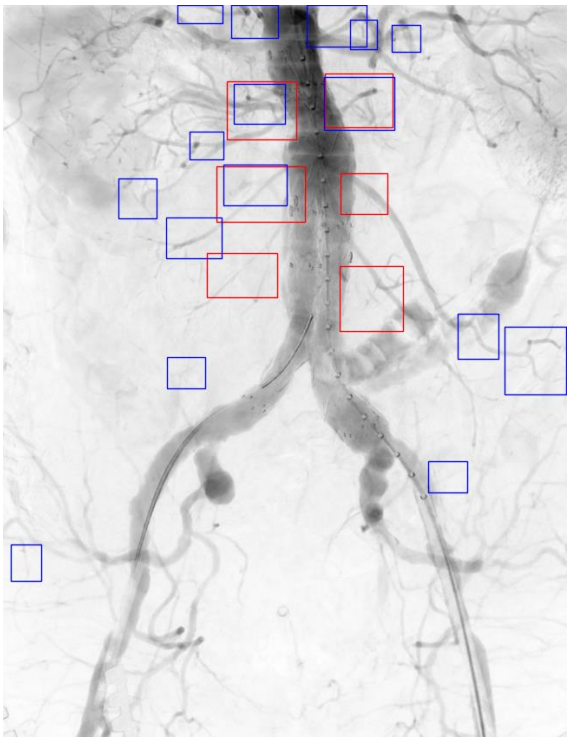

Fig. 19. Illustration of a prediction at a low threshold for creating the binary prediction. The superior mesenteric artery (SMA) and branches of the SMA are accountable for a large number of FP predictions. The red boxes represent the ground truth, the blue boxes represent the prediction of the network.

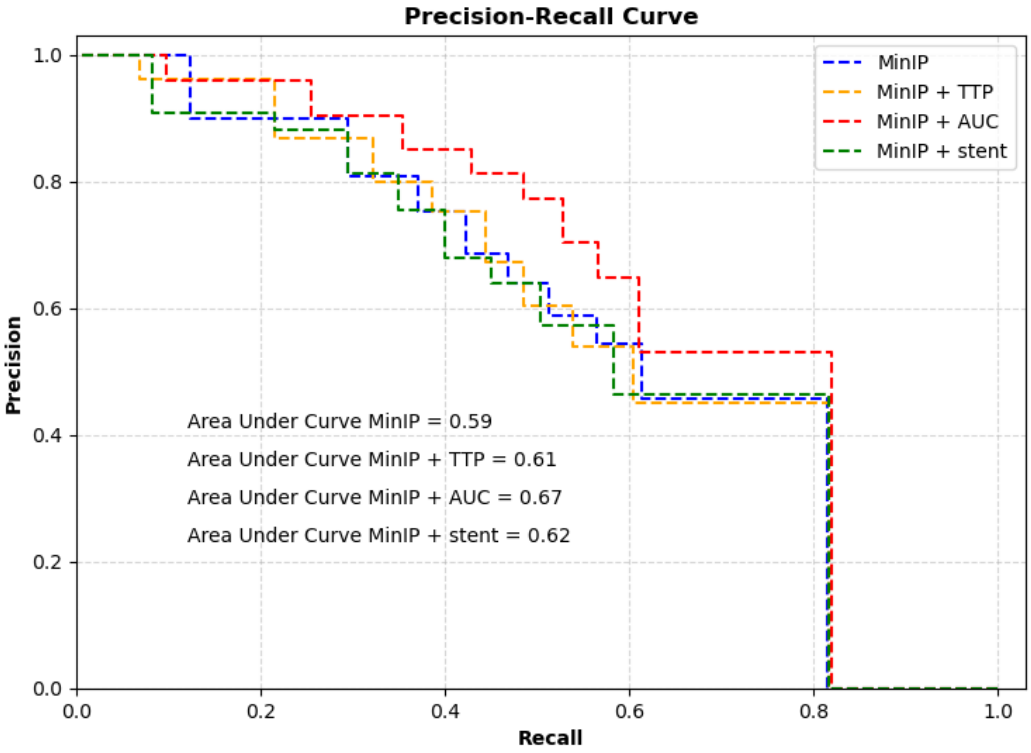


Fig. 20. Precision-recall curves of the U-Net trained on different combinations of input projections. The precision-recall curve is determined based on the complete cross-validation of 70 cDSAs for each of the combinations of input projections.

## 4.4 Discussion

This study has demonstrated that a U-Net can be utilized to automatically identify and localize lumbar arteries on a completion angiogram, based on 2D feature projections obtained from the cDSA serie. In light of the automatic detection and classification of endoleaks on a completion angiogram using deep learning, this trained network can aid in the classification of endoleaks after incorporation to a network for automatic detection of endoleaks. The added value for classification of endoleaks primarily applies to the classification of type 2 endoleaks due to the involvement of lumbar arteries for this type of endoleak.

### *4.4.1 Interpretation of the results*

Analysis of the qualitative results have shown that the trained U-Net is capable of correctly identifying and localizing lumbar arteries on a completion angiogram. The results, as in Fig. 18 and Fig. 19, also show that post-processing of the prediction heatmaps returned by the network plays an important role in the performance of the final prediction of the network. The majority of the false positive predictions are the result of activation in the prediction heatmap on the contrast-filled superior mesenteric artery branches. The incorrect activations on the branches of the superior mesenteric artery may the be result of their close proximity to the lumbar arteries on the completion angiogram and their relatively similar diameter. Concerning the diameter of the artery, there were no activations in the prediction heatmap at the site of the renal arteries, indicating that the diameter may play a role in distinguishing between different arteries.

With the reduction of the cDSA series into a single 2D MinIP projection, temporal information is lost. Therefore, the hypothesis was that extra 2D feature projections added to the MinIP projection would increase the performance of the network. The quantitative results show that the network with only the MinIP projection as input resulted in the lowest $AUC_{PR}$. The addition of extra 2D feature projections increased the performance of the network. The AUC projection, which is a measure for the amount of contrast within a pixel throughout the cDSA serie, has the most added value to the performance of the network. The accuracy of the stent-graft segmentation mask may influence the added value of the stent-graft segmentation mask as an additional input projection. Visual inspection, however, did not show a clear relation between the accuracy of the stent-graft and the prediction heatmap in terms of correct identification and localization of the lumbar arteries.

The false negative predictions can partly be explained based on the visibility of the lumbar arteries on the 2D projections that were used as input. During annotation of the lumbar arteries, the contrast levels of the cDSAs were adjusted for better visualization of the arteries with limited contrast-agent and visibility. Lumbar arteries with poor visibility, but visible with adjustment of contrast levels, were also annotated. The visibility of these barely visible lumbar arteries on the cDSA was lost during creation of the 2D projections, and may therefore be a plausible cause of the false negative predictions. An important finding

is that part of the false negative predictions are thus a result of the loss of visibility of the lumbar arteries in the input data to the network and not a result of post-processing.

### *4.4.2 Limitations of the study*

The primary limitation of the current study is the limited availability of cDSAs for training the U-Net for automatic identification of lumbar arteries. There is a significant variability among the cDSA images acquired at the end of an EVAR procedure between different patients in terms of pixel intensities, which are dependent on the body composition of patients and influences the visibility of the lumbar arteries on the cDSA. Furthermore, the amount, location and course of lumbar arteries can vary significantly between patients. The data set needs to be extended to account for these differences across the patient's anatomy and is expected to increase the performance of the network.

The performance of the trained U-Net has not been validated on an external data set. The cDSA images in the data set are all acquired in the same type of hybrid operating room in a single institution. Therefore, no conclusions can be drawn about the robustness of the network in other institutions and on cDSAs acquired using hybrid operating rooms of different manufacturers.

The identification of the lumbar arteries on the cDSA was faced with a 2D approach, using the 2D feature projections obtained from the cDSA. While this approach has its benefits as described in section 4.2.3, this approach also comes with a limitation. Reduction of the 2D + time cDSA series to a single frame also results in possible over-projection of lumbar arteries. This may limit the ability of the deep learning network to correctly detect the lumbar arteries. The extent of over-projection of lumbar arteries by other arteries has not been studied in this thesis, and should be considered for further research on this topic.

### *4.4.3 Future perspectives*

In future studies, there are several aspects that could be focused on to further improve the performance of the U-Net for identification of lumbar arteries. First of all, this study has evaluated the performance of the network based on the MinIP projection and with the addition of extra 2D feature projections. Only a single extra 2D feature projection was added along with the MinIP projection in this study. In future work, the influence of adding multiple 2D feature projections along with the MinIP projection on the performance of the network can be investigated.

Another aspect that could be focused on are the post-processing steps to obtain a binary prediction from the prediction heatmaps. In this study, the initial binary prediction from the prediction heatmap was performed using a threshold for pixel intensity, and subsequently, components with an area below a fixed threshold were removed as they were considered artefacts. Within the post-processing steps, only the pixel intensity threshold was varied while other parameters were fixed. In future work, the influence of other parameters in the post-processing steps on the final prediction and performance of the network could be investigated.

After further optimization of the current methods, the network for automatic identification and localization of lumbar arteries should be incorporated into a network for the automatic detection and classification of endoleaks on cDSA images. The current network can be of added value for classification of, especially, type 2 endoleaks due to their close relation with lumbar arteries. In such a future network, the performance of the classification of type 2 endoleaks can be investigated based on the close proximity or overlap between the detected endoleak and lumbar arteries.

## 4.5 Conclusion

This study has demonstrated the usability of a U-Net for the automatic identification and localization of the lumbar arteries on completion digital subtraction angiographies acquired at the end of an EVAR procedure. The addition of extra 2D feature projections increased the performance of the network, with the AUC projection resulting in the largest increase of performance.

Chapter 5

# General discussion for clinical implementation and overall conclusion

## 5.1 General discussion for clinical implementation

This thesis has demonstrated that deep learning can be utilized to automatically extract information from completion digital subtraction angiography images. More specifically, we have shown the possibility of automatic detection of renal arteries and lumbar arteries, as part of the ultimate goal of objective intraoperative image analysis during EVAR procedures. This section describes the usability and clinical implementation of these developed deep learning-based methods.

Digital subtraction angiography has the limitation of image acquisition in a single direction. This technique is thus prone to over-projection, possibly leading to the invisibility of lumbar arteries due to coverage of the arteries by the contrast-filled aortic lumen. This could limit the detection of relevant lumbar arteries by a deep learning network. However, in this thesis, we have demonstrated that only around two percent of the lumbar arteries are over-projected by the aortic lumen. Therefore, the uni-directional image acquisition does not limit the automatic detection of lumbar arteries using deep learning models in terms of over-projection.

Invisibility of a lumbar artery on the CTA has a good predictive value for the invisibility of the same lumbar artery on the completion angiogram. In contrast, the visibility of the lumbar artery on the CTA has a moderate predictive value for the visibility of the same lumbar artery on the completion angiogram. Further research could focus on the cause of the reduced predictive value for the latter case. A good predictive value for the visibility of lumbar arteries on the completion angiogram could function as a form of prior knowledge in the automatic detection of lumbar arteries.

Given further optimization of the renal artery ostia detection network, this model can aid in the objective analysis of the stent-graft position analysis. During deployment of the stent-graft in infrarenal aneurysms with a short neck, a minimum apposition needs to be achieved to ensure a proper seal of the endoprosthesis. A combination of an optimized renal artery ostia network and stent-graft segmentation network can be utilized to this end. Infrarenal or suprarenal fixation of the stent-graft needs to be taken into account when developing the deployment accuracy analysis.

As described in the chapter 4, the renal artery ostia detection network can also be utilized as a predictor for classification of persistent and transient endoleaks in case of early onset type 2 endoleaks. The basis of this analysis lays in the study published by Mursalin et al., who described a significant difference between the appearance time of the endoleak cavity after contrast-agent filling of the renal arteries between persistent and transient endoleaks with a sensitivity of 88% and specificity of 86%.[25] To achieve this automatic analysis, a combination of the renal artery ostia detection network and endoleak detection network needs to be realised.

Automatic identification of lumbar arteries on the final completion angiogram can be of added value for the automatic classification of endoleaks, especially type 2 endoleaks. As the treatment options and treatment urgency for the different type of endoleaks differ significantly, the automatic classification of

endoleaks can impact the clinical decision-making on whether or not further surgical acting is required within the same EVAR procedure.

## 5.2 Overall conclusion

The overall aim of this study was to develop automatic methods for the detection and localization of renal artery ostia and lumbar arteries on the completion digital subtraction angiography at the end of an EVAR procedure. The future goal is to use these methods for further development of objective intraoperative analytical applications, such as stent-graft position evaluation and endoleak classification.

In the first phase of this study, we have demonstrated that lumbar arteries are rarely over-projected by the aortic lumen on completion angiograms, eliminating over-projection as an obstacle for using completion angiograms for automatic lumbar artery identification. Furthermore, we have concluded that the invisibility of a lumbar artery on the pre-operative CTA is a good predictor for the invisibility of the same lumbar artery on the intraoperative cDSA, while a visible lumbar artery on the CTA is a moderate predictor for a visible lumbar artery on the completion angiogram.

In the second phase of the study, we demonstrated the possibility of automatic localization of the renal artery ostia on the cDSA using a deep learning network. Using the proposed methods, three quarters of the renal artery ostia were automatically detected within a clinically relevant range. Automatic renal artery detection can aid in objective stent-graft position analysis and sub-classification of type 2 endoleaks.

In the third phase of the study, we demonstrated the feasibility of a deep learning based method for the automatic identification of lumbar arteries on the final completion angiogram. Automatic identification of lumbar arteries can be of added value for automatic classification of type 2 endoleaks.

# References


1. Sakalihasan, N. *et al.* Abdominal aortic aneurysms. *Nat. Rev. Dis. Prim.* **4**, (2018).
2. Wanhainen, A. *et al.* European Society for Vascular Surgery (ESVS) 2019 Clinical Practice Guidelines on the Management of Abdominal Aorto-iliac Artery Aneurysms. (2018) doi:10.1016/j.ejvs.2018.09.020.
3. Chaikof, E. L. *et al.* The Society for Vascular Surgery practice guidelines on the care of patients with an abdominal aortic aneurysm. *J. Vasc. Surg.* **67**, 2-77.e2 (2018).
4. Brewster, D. C. *et al.* Guidelines for the treatment of abdominal aortic aneurysms: Report of a subcommittee of the Joint Council of the American Association for Vascular Surgery and Society for Vascular Surgery. *J. Vasc. Surg.* **37**, 1106–1117 (2003).
5. Zommorodi, S., Bottai, M. & Hultgren, R. Sex differences in repair rates and outcomes of patients with ruptured abdominal aortic aneurysm. *Br. J. Surg.* **106**, 1480–1487 (2019).
6. Dutch Surgical Aneurysm Audit Jaarrapportage 2018. https://dica.nl/jaarrapportage-2018/dsaa (2018).
7. Breininger, K. *et al.* Intraoperative stent segmentation in X-ray fluoroscopy for endovascular aortic repair. *Int. J. Comput. Assist. Radiol. Surg.* **13**, 1221–1231 (2018).
8. Prinssen, M. *et al.* A Randomized Trial Comparing Conventional and Endovascular Repair of Abdominal Aortic Aneurysms. *N. Engl. J. Med.* **351**, 1607–1618 (2004).
9. Demehri, S. *et al.* Volumetric quantification of type II endoleaks: An indicator for aneurysm sac growth following endovascular abdominal aortic aneurysm repair. *Radiology* **271**, 282–290 (2014).
10. Swerdlow, N. J., Wu, W. W. & Schermerhorn, M. L. Open and endovascular management of aortic aneurysms. *Circ. Res.* **124**, 647–661 (2019).
11. Greenhalgh, R. M. Comparison of endovascular aneurysm repair with open repair in patients with abdominal aortic aneurysm (EVAR trial 1), 30-day operative mortality results: Randomised controlled trial. *Lancet* **364**, 843–848 (2004).
12. De Bruin, J. L. *et al.* Long-Term Outcome of Open or Endovascular Repair of Abdominal Aortic Aneurysm. *N. Engl. J. Med.* **362**, 1881–1889 (2010).
13. van Schaik, T. G. *et al.* Long-term survival and secondary procedures after open or endovascular repair of abdominal aortic aneurysms. in *Journal of Vascular Surgery* vol. 66 1379–1389 (Mosby Inc., 2017).
14. Powell, J. T. *et al.* Meta-analysis of individual-patient data from EVAR-1, DREAM, OVER and ACE trials comparing outcomes of endovascular or open repair for abdominal aortic aneurysm over 5 years. *British Journal of Surgery* vol. 104 166–178 (2017).
15. Orgera, G. *et al.* Techniques and future perspectives for the prevention and treatment of endoleaks after endovascular repair of abdominal aortic aneurysms. *Insights Imaging* **10**, (2019).
16. White, S. B. & Stavropoulos, S. W. Management of endoleaks following endovascular aneurysm repair. *Semin. Intervent. Radiol.* **26**, 33–38 (2009).
17. Lal, B. K. *et al.* Predictors and outcomes of endoleaks in the Veterans Affairs Open Versus Endovascular Repair (OVER) Trial of Abdominal Aortic Aneurysms. in *Journal of Vascular Surgery* vol. 62 1394–1404 (Mosby Inc., 2015).
18. Chen, J. & Stavropoulos, S. W. Management of Endoleaks. *Semin. Intervent. Radiol.* **32**, 259–264 (2015).

19. Swerdlow, N. J. *et al.* Select type I and type III endoleaks at the completion of fenestrated endovascular aneurysm repair resolve spontaneously. *J. Vasc. Surg.* **70**, 381–390 (2019).

20. Bashir, M. R., Ferral, H., Jacobs, C., McCarthy, W. & Goldin, M. Endoleaks after endovascular abdominal aortic aneurysm repair: Management strategies according to CT findings. *Am. J. Roentgenol.* **192**, (2009).

21. Lee, J. T. *et al.* EVAR Deployment in Anatomically Challenging Necks Outside the IFU 5. (2013) doi:10.1016/j.ejvs.2013.03.027.

22. Madigan, M. C., Singh, M. J., Chaer, R. A., Al-Khoury, G. E. & Makaroun, M. S. From the Vascular and Endovascular Surgery Society Occult type I or III endoleaks are a common cause of failure of type II endoleak treatment after endovascular aortic repair. (2019) doi:10.1016/j.jvs.2018.04.054.

23. Müller-Wille, R. *et al.* CT features of early type II endoleaks after endovascular repair of abdominal aortic aneurysms help predict aneurysm sac enlargement. *Radiology* vol. 274 906–916 (2015).

24. Quinn, A. A. *et al.* The incidence and fate of endoleaks vary between ruptured and elective endovascular abdominal aortic aneurysm repair. in *Journal of Vascular Surgery* vol. 65 1617–1624 (Mosby Inc., 2017).

25. Mursalin, R. *et al.* Imaging-based predictors of persistent Type II endoleak after endovascular abdominal aortic aneurysm repair. *Am. J. Roentgenol.* **206**, 1335–1340 (2016).

26. Arenas Azofra, E. *et al.* Predictive Factors of Aneurysm Sac Growth in Patients with a Type II Endoleak in the First Post-EVAR Control. *Ann. Vasc. Surg.* **68**, 245–251 (2020).

27. Maleux, G. *et al.* Incidence, etiology, and management of type III endoleak after endovascular aortic repair. *J Vasc Surg* **66**, 1056–64 (2017).

28. UpToDate. Types of endoleak after endovascular repair - UpToDate. https://www.uptodate.com/contents/image?imageKey=CARD%2F64707&topicKey=SURG%2F15221&source=see_link (2022).

29. Schlösser, F. J. V. *et al.* Aneurysm Rupture after EVAR: Can the Ultimate Failure be Predicted? *Eur. J. Vasc. Endovasc. Surg.* **37**, 15–22 (2009).

30. Sidloff, D. A., Stather, P. W., Choke, E., Bown, M. J. & Sayers, R. D. Type II endoleak after endovascular aneurysm repair. *British Journal of Surgery* vol. 100 1262–1270 (2013).

31. Lo, R. C. *et al.* Risk factors and consequences of persistent type II endoleaks. in *Journal of Vascular Surgery* vol. 63 895–901 (Mosby Inc., 2016).

32. Antoniou, G. A., Alfahad, A., Antoniou, S. A. & Torella, F. Prognostic Significance of Aneurysm Sac Shrinkage After Endovascular Aneurysm Repair. *J. Endovasc. Ther.* **27**, 857–868 (2020).

33. Jebbink, G. *et al.* Citation: van Rijswijk Clinical Medicine Predictors of Abdominal Aortic Aneurysm Shrinkage after Endovascular Repair. *J. Clin. Med* **2022**, 1394 (2022).

34. Couchet, G. *et al.* Predictive Factors for Type II Endoleaks after Treatment of Abdominal Aortic Aneurysm by Conventional Endovascular Aneurysm Repair. *Annals of Vascular Surgery* vol. 29 1673–1679 (2015).

35. Seike, Y. *et al.* The Influence of 4 or more Patent Lumbar Arteries on Persistent Type II Endoleak and Sac Expansion after Endovascular Aneurysm Repair. *Ann. Vasc. Surg.* **50**, 195–201 (2018).

36. Lau, C., Feldman, D. N., Girardi, L. & Kim, L. Imaging for surveillance and operative management for endovascular aortic aneurysm repairs. *J. Thorac. Dis.* **9**, S309–S316 (2017).

37. Wacker, F. K., Valdeig, S., Raatschen, H. J. & Meyer, B. C. C-Arm CT - An adjunct to DSA for endoleak classification in patients with endovascular repair of abdominal aortic aneurysms. *RoFo Fortschritte auf dem Gebiet der Rontgenstrahlen und der Bildgeb. Verfahren* **186**, 247–252 (2014).

38. Ward, T. J. *et al.* Anatomic risk factors for type-2 endoleak following EVAR: a retrospective review of preoperative CT angiography in 326 patients. *Cardiovasc. Intervent. Radiol.* **37**, 324–328 (2014).

39. Greenhalgh, R. Comparison of endovascular aneurysm repair with open repair in patients with abdominal aortic aneurysm (EVAR trial 1), 30-day operative mortality results: Randomised controlled trial. *Lancet* **364**, 843–848 (2004).

40. De Bruin, J. L. *et al. Long-Term Outcome of Open or Endovascular Repair of Abdominal Aortic Aneurysm A bs tr ac t*.

41. Gezondheidsraad. *Incidentie, mortaliteit en prevalantie AAA in Nederland*. (2019).

42. Ronneberger, O., Fischer, P. & Brox, T. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. http://lmb.informatik.uni-freiburg.de/.

43. Payer, C., Štern, D., Bischof, H. & Urschler, M. Regressing Heatmaps for Multiple Landmark Localization Using CNNs. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* **9901 LNCS**, 230–238 (2016).

44. Cina, A. *et al.* 2-Step Deep Learning Model for Landmarks Localization in Spine Radiographs. *Sci. Rep.* **11**, 1–12 (2021).

45. MeVisLab - Image Processing Research and Development. (2022).

46. Consortium, M. MONAI: Medical Open Network for AI. (2020) doi:10.5281/ZENODO.6114127.

47. Noothout, J. M. H. *et al.* Deep Learning-Based Regression and Classification for Automatic Landmark Localization in Medical Images. *IEEE Trans. Med. Imaging* **39**, 4011–4022 (2020).

48. Sidloff, D. A. Editor ' s Choice e Type II Endoleak : Conservative Management Is a Safe Strategy. *Eur. J. Vasc. Endovasc. Surg.* **48**, 391–399 (2013).

49. Oderich, G. S. *et al.* Reporting standards for endovascular aortic repair of aneurysms involving the renal-mesenteric arteries. *J. Vasc. Surg.* **73**, 4S-52S (2020).

# Appendix A

## Deep learning-based intraoperative stent-graft segmentation on completion digital subtraction angiography during Endovascular Aneurysm Repair

*Technical Note*

*Submitted*

# Deep Learning-Based Intraoperative Stent-Graft Segmentation on Completion Digital Subtraction Angiography during Endovascular Aneurysm Repair

Kaj O. Kappe, BSc.*[1,2], Stefan P.M. Smorenburg, MSc.*[1], Arjan W.J. Hoksbergen, MD, PhD[1], Jelmer M. Wolterink, PhD[3], Kak Khee Yeung, MD, PhD, FEBVS[1]

[1] *Department of Surgery, Amsterdam Cardiovascular Sciences, Amsterdam University Medical Centers, Vrije Universiteit, Amsterdam, the Netherlands*

[2] *Technical Medicine, Faculty of Science and Technology, University of Twente, Enschede, the Netherlands*

[3] *Mathematics of Imaging & AI Group, Faculty of Electrical Engineering, Mathematics, and Computer Science, University of Twente, Enschede, the Netherlands*

* *Kaj O. Kappe and Stefan P.M. Smorenburg contributed equally.*

**Abstract**

**Purpose:**

Modern endovascular hybrid operating rooms generate large amounts of medical images during a procedure, which are currently mostly assessed by eye. In this paper, we present fully automatic segmentation of the stent graft in completion digital subtraction angiography during endovascular aneurysm repair, utilizing a deep learning network.

**Technique:**

Completion digital subtraction angiographies (cDSAs) of 47 patients treated for an infrarenal aortic aneurysm using EVAR were collected retrospectively. A two-dimensional convolutional neural network (CNN) with a U-Net architecture was trained for segmentation of the stent graft from the completion angiographies. The cross-validation resulted in an average Dice similarity score of 0.957 ± 0.041 and median of 0.968 (IQR: 0.950 – 0.976). The mean and median of the average surface distance are 1.266 ± 1.506 mm and 0.870 mm (IQR: 0.490 – 1.430), respectively.

**Conclusion:**

We developed a fully automatic stent graft segmentation method based on the completion digital subtraction angiography during EVAR, utilizing a deep learning network. This can provide the platform for the development of intraoperative analytical applications in the endovascular hybrid operating room such as stent graft deployment accuracy, endoleak visualization, and image fusion correction.

**Keywords:**

Deep learning, Artificial Intelligence, Digital Subtraction Angiography, Endovascular Aneurysm Repair, EVAR, Segmentation, Automatic, Intraoperative, Abdominal Aortic Aneurysm, AAA

## Introduction

Since its introduction, endovascular aneurysm repair (EVAR) has been inextricably dependent on intraoperative fluoroscopic imaging. Technical innovations in the past decade have led to the replacement of the mobile fluoroscopic C-arm with modern high-tech hybrid operating rooms. With the increased imaging capabilities of the contemporary hybrid operating room, large numbers of fluoroscopy images and digital subtraction angiography (DSA) images are generated during and after the EVAR procedure. Intraoperative clinical decision making has been predominantly based on visual inspection of images by the operating team, with the exception of CT-fluoroscopy image fusion for navigational purposes[1,2]. However, while hundreds of images are acquired during a typical intervention, the vast majority of these go unused, while they may have significant value to add information to the procedure or improve procedural outcomes.

The interest in artificial intelligence techniques, and particularly deep learning, has exploded in the past years, driven by an increase in computational power, the utilization of large datasets, and image-guided surgery.[3] We propose that these techniques can be used to fully exploit all acquired intraoperative images at no additional burden to the operating team. A considerable advantage of using deep learning algorithms is the objective analysis  images versus the subjective visual inspection by the operating team. Moreover, because the number of examples that a deep learning algorithm is exposed to during training is only limited by the availability of data, its exposure could be higher than that of a single observer, potentially outperforming humans. In terms of clinical implementation, this has the potential of improving interobserver agreement during clinical decision-making in the operating room.

Completion DSAs (cDSAs) performed at the end of the procedure after stent graft deployment contain information on the stent graft's position, possible endoleaks, patency of arteries and stent-graft limbs, and blood flow dynamics. To fully exploit the possible hidden and valuable information in these images to improve procedural outcomes, we propose to analyse these images using deep learning. In this paper, we present fully automatic segmentation of the stent-graft in cDSA, utilizing a deep learning network. We hypothesize that an automatic stent graft segmentation method can be developed which can provide the basis for further intraoperative applications to objectively analyse the cDSAs in the endovascular hybrid operation room.

## Technique

### *Data collection*

Completion digital subtraction angiographies of 47 patients treated for an infrarenal aortic aneurysm using EVAR were collected retrospectively. Patients in the study cohort received a bifurcated stent graft, bifurcated stent graft with an iliac branch device, or a single-limb aortoiliac stent graft. The collection of

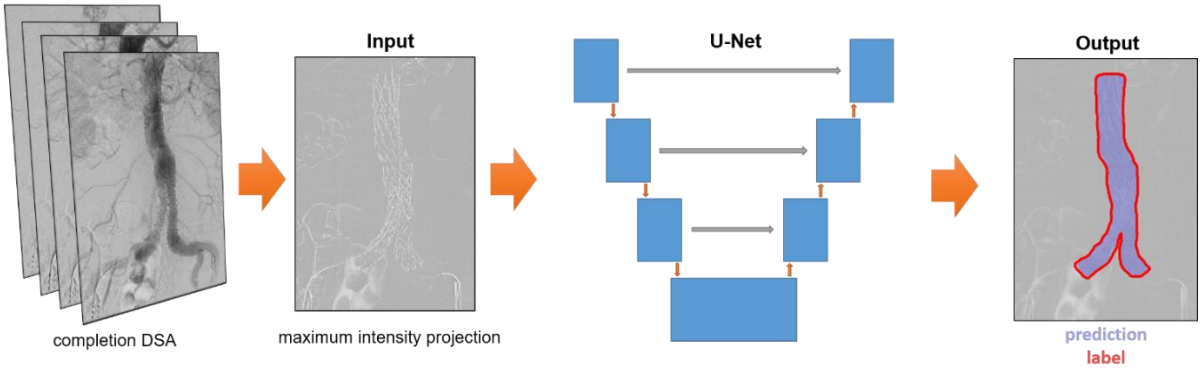


**Figure 1.** Overview of machine learning network with minimum intensity projection and label as input and corresponding output. The output image has a Dice score of 0.979.

cDSAs from the patients' electronic medical records was performed with the approval of the local medical ethical review board. All cDSAs were acquired using a C-arm X-ray system in a Philips Azurion FlexMove 7 C20 Hybrid Operating Room (Philips Healthcare, Best, The Netherlands) in the Amsterdam University Medical Centers, location VUmc, in the period from May 2017 to April 2020. Images were acquired with isotropic and anisotropic dimensions ranging from 790 x 1024 to 1904 x 1904 pixels with isotropic pixel sizes ranging from 0.154 x 0.154 $mm^2$ to 0.370 x 0.370 $mm^2$. cDSAs were acquired with a multiphase acquisition protocol; phase 1 for 6 seconds at 3 frames per second (fps), phase 2 for 5 seconds at 1 fps and phase 3 at 0.5 fps indefinitely. Typically one cDSA contained between 18 to 43 images per series. A purposely induced apnoea minimalized visceral and stent graft motion during the recording of all series.

### *Image modification*

We assumed that the stent graft did not move during a cDSA series and therefore posed our segmentation problem as a two-dimensional (2D) segmentation problem, in which one segmentation stent graft mask is obtained per cDSA series. To represent a cDSA series as a 2D image, a maximum intensity projection (MaxIP) was obtained along its temporal axis. This resulted in one projection image per cDSA, in which the visibility of the injected contrast agent was minimized, thereby maximizing the visibility of the stent graft. To train and evaluate the deep learning method, pixel-wise manual segmentation masks of the stent graft were created by an expert for all cDSA series. Subsequently, all 2D projection images and corresponding ground truth masks were resampled to 0.4 x 0.4 $mm^2$ isotropic pixel spacing using bilinear interpolation and nearest neighbour interpolation, respectively.

### *Neural network architecture and training*

A two-dimensional convolutional neural network (CNN) with a U-Net architecture was trained for segmentation of the stent graft from the 2D projection images.[12] Training of the neural network was performed using randomly extracted images patches with size 512 x 512 pixels from the projection images. Training patches were augmented to create additional synthetic data to increase the

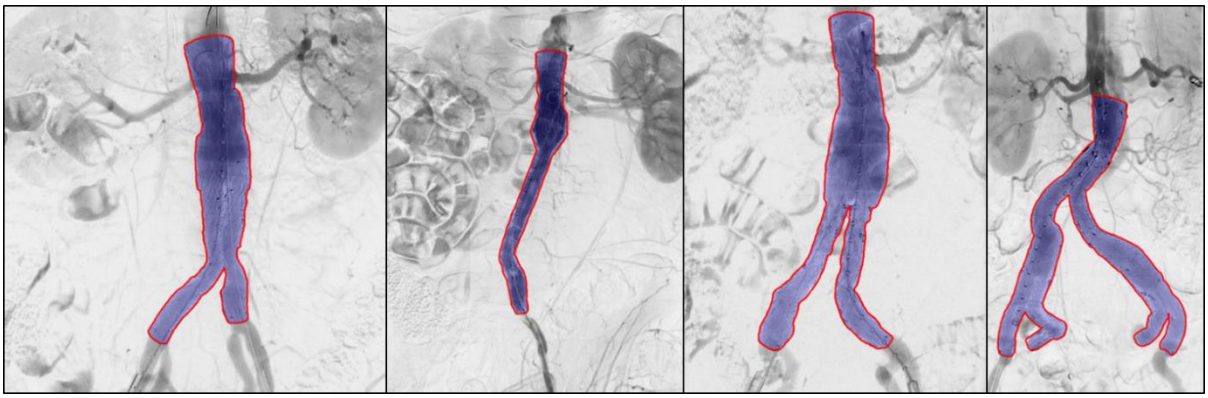

**Figure 2.** Successful automatic stent segmentation results (purple) and corresponding ground truth labels (red). Note that these the first three stent grafts have suprarenal fixation and the prediction is based on the stent graft struts and not the fabric. The Dice coefficient scores for the shown predictions are 0.985, 0.976, 0.983 and 0.976

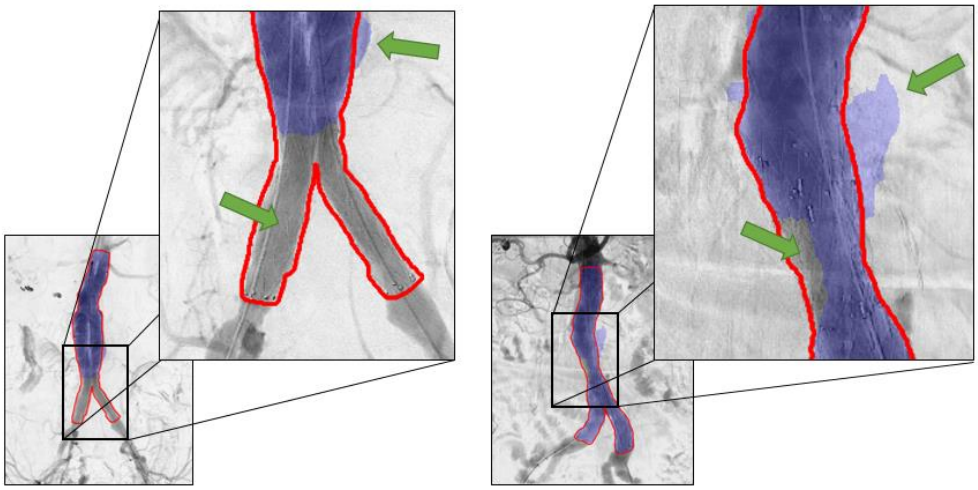

**Figure 3.** Two patients with suboptimal segmentation. The upper green arrows indicate an area predicted outside the ground truth label. The lower green arrows indicate an area not predicted inside the ground truth label. The left prediction has a Dice score of 0.719 and the right prediction has a Dice score of 0.930

generalizability of the network. Data augmentation consisted of horitzontal flipping with a probability of 0.5 and rotation around the image centre with a maximum deviation of ± 14 degrees. The network was trained by minimizing a Dice loss with an Adam optimizer algorithm using a learning rate of 0.001, dropout of 0.2, and a batch size of 8. The training was performed for 2000 epochs with a reduction of the learning rate by a factor 10 after 1000 and 1500 epochs.

The network was implemented using the Pytorch-based framework Medical Open Network for AI (MONAI)[4] and run on a workstation with a NVIDIA GeForce RTX 3090 Graphics Card with 24 GB of memory.

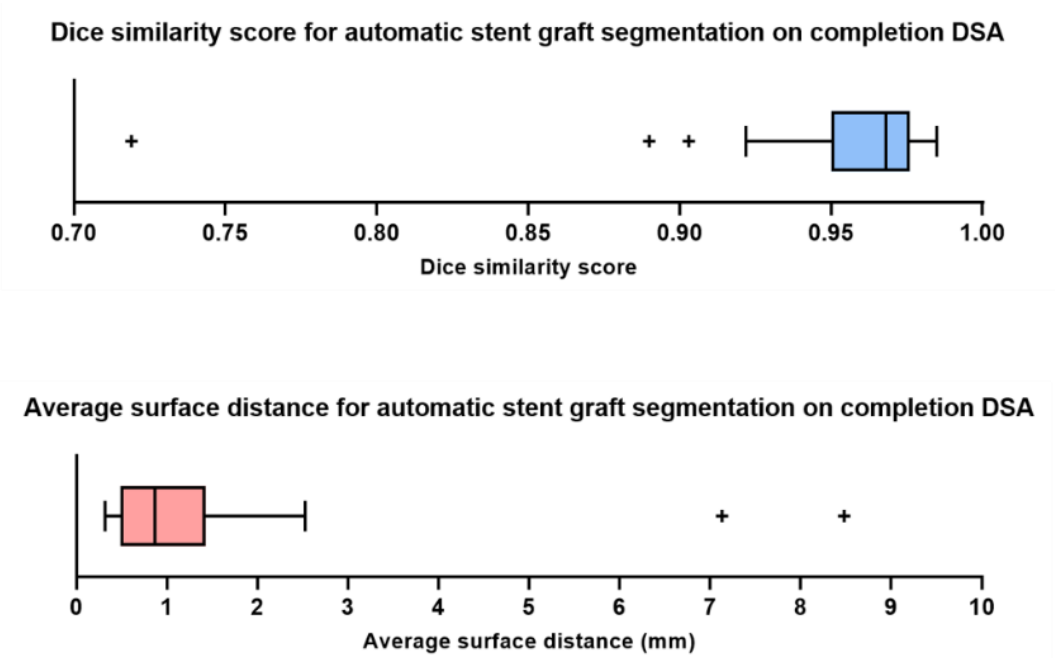


**Figure 4.** Box plots demonstrating the quantitative results (Dice coefficient score and average surface distance) of the cross-validation on then full data set.

### *Evaluation of the network's performance*

The performance of the network was evaluated based on a Dice similarity score, which measures the overlap of the predicted segmentation of the model with the ground truth segmentation. Moreover, we computed the average surface distance between the reference and predicted stent graft masks. This is the average distance (in mm) from all points on the boundary of the predicted segmentation to the closest point on the boundary of the ground truth segmentation.

We performed nine-fold cross-validation on the full data set. Automatic stent graft segmentations with a high Dice similarity score on unseen test images are shown in Fig. 2. The results show that the network is capable of segmenting bifurcated stent-grafts, aorto-uni-iliac stent grafts, and bifurcated stent grafts with iliac branch devices, irresepective of the device manufacturer. Moreover, the network is able to correctly annotate stent grafts with infrarenal and suprarenal fixation. Fig. 3. shows the result of the largest outlier in Fig. 4, together with another suboptimal segmentation as a result of over- and under-segmentation.

Quantitative results are shown in Figure 4. A Shapiro-Wilk test and Q-Q plots show that both the Dice similarity scores and average surface distances are non-normally distributed. The cross-validation resulted in an average Dice similarity score of 0.957 ± 0.041 and median of 0.968 (IQR: 0.950 – 0.976). The mean and median of the average surface distance are 1.266 ± 1.506 mm and 0.870 mm (IQR: 0.490 – 1.430), respectively.

## Discussion

The results of our study demonstrate that a deep learning-based method can fully automatically segment the stent graft in completion digital subtraction angiography images. The proposed method can provide the basis for further development of clinical analytic applications of cDSAs. This includes objective assessment of stent graft deployment, present endoleaks, patency of arteries and stent graft limbs and blood flow dynamics.

Inspection of the results shows that our proposed method is capable of segmenting different configurations of stent grafts, i.e. bifurcated stent grafts with and without iliac branch devices and aorto-uni-iliac stent grafts. The different strut structures and suprarenal and infrarenal fixation did not seem to influence the performance of the deep learning network.

Several previous studies have reported automatic feature extraction of angiographic images during EVAR. These studies mainly focus on characterization and quantification of arterial deformations after insertion of stiff guide wires, intraoperative correction of image fusion and automatic segmentation of guidewires.[5-10] Closest to our work is the study of Breininger et al., who also proposed a fully automatic deep learning based method for stent graft segmentation in single 2D fluoroscopic images.[11] Note that in contrast to their method, we perform automatic segmentation on a 2D projection obtained from the entire 2D + time cDSA sequence. This leads to results that appear to be slightly more accurate than those reported by Breininger et al., who obtained an average Dice coefficient score of $0.943 \pm 0.043$ whereas the average of our Dice coefficient was $0.957 \pm 0.041$.

The study was limited by the number of available cDSAs. Data augmentation was utilized to partially compensate for the reduced number of available imaging. However, expansion of the dataset would likely further improve the performance and robustness of the deep learning network. Despite the high Dice similarity scores of our experiment, one major outlier as a result of under segmentation can be observed in Fig. 4 of which the result is visualized in Fig. 3. Limited visibility and contrast of the stent graft limbs on the 2D projection may be the underlying cause. Over-segmentation, shown in Fig. 3 as well, may be due to substantial background artefacts due to existing contrast-agent in the intestines. Additional input channels to the network comprising different 2D projections of the cDSA may help reduce over- and under segmentation.

Our proposed method for fully automatic stent graft segmentation lays the foundation for further development of clinical analytic applications for cDSA to fully exploit possible valuable information. Future work can focus on the development of deep learning networks for segmentation of anatomical structures and analysis of flow parameters. The merging of separate models can lead to the development of fully automatic analytic applications to assist the surgeon in more objective clinical decision making in the hybrid operation room.

## Conclusion

We developed a fully automatic stent graft segmentation method based on the completion digital subtraction angiography during EVAR, utilizing a deep learning network. This can provide the platform for the development of intraoperative analytical applications in the endovascular hybrid operating room such as stent graft deployment accuracy, endoleak information, and image fusion correction.

## Disclosures

This research is supported by an unrestricted research grant of Philips (Best, the Netherlands).

## References


1. Doelare SAN, Smorenburg SPM, van Schaik TG, Blankensteijn JD, Wisselink W, Nederhoed JH, et al. Image Fusion During Standard and Complex Endovascular Aortic Repair, to Fuse or Not to Fuse? A Meta-analysis and Additional Data From a Single-Center Retrospective Cohort. Journal of Endovascular Therapy. 2020.
2. Stangenberg L, Shuja F, Carelsen B, Elenbaas T, Wyers MC, Schermerhorn ML. A novel tool for three-dimensional roadmapping reduces radiation exposure and contrast agent dose in complex endovascular interventions. J Vasc Surg. 2015;62(2):448-55.
3. Soffer S, Ben-Cohen A, Shimon O, Amitai MM, Greenspan H, Klang E. Convolutional Neural Networks for Radiologic Images: A Radiologist's Guide. Radiology. 2019;290(3):590-606.
4. (2020) MC. MONAI: Medical Open Network for AI (0.6.0). [
5. Breininger K, Hanika M, Weule M, Kowarschik M, Pfister M, Maier A. Simultaneous reconstruction of multiple stiff wires from a single X-ray projection for endovascular aortic repair. Int J Comput Assist Radiol Surg. 2019;14(11):1891-9.
6. Cercenelli L, Bortolani B, Tiberi G, Mascoli C, Corazza I, Gargiulo M, et al. Characterization of Vessel Deformations During EVAR: A Preliminary Retrospective Analysis to Improve Fidelity of Endovascular Simulators. J Surg Educ. 2018;75(4):1096-105.
7. Gindre J, Bel-Brunon A, Rochette M, Lucas A, Kaladji A, Haigron P, et al. Patient-Specific Finite-Element Simulation of the Insertion of Guidewire During an EVAR Procedure: Guidewire Position Prediction Validation on 28 Cases. IEEE Trans Biomed Eng. 2017;64(5):1057-66.
8. Kaladji A, Dumenil A, Castro M, Cardon A, Becquemin JP, Bou-Said B, et al. Prediction of deformations during endovascular aortic aneurysm repair using finite element simulation. Comput Med Imaging Graph. 2013;37(2):142-9.
9. Lessard S, Kauffmann C, Pfister M, Cloutier G, Therasse E, de Guise JA, et al. Automatic detection of selective arterial devices for advanced visualization during abdominal aortic aneurysm endovascular repair. Med Eng Phys. 2015;37(10):979-86.
10. Mohammadi S, Mohammadi M, Dehlaghi V, Ahmadi A. Automatic Segmentation, Detection, and Diagnosis of Abdominal Aortic Aneurysm (AAA) Using Convolutional Neural Networks and Hough Circles Algorithm. Cardiovasc Eng Technol. 2019;10(3):490-9.
11. Breininger K, Albarqouni S, Kurzendorfer T, Pfister M, Kowarschik M, Maier A. Intraoperative stent segmentation in X-ray fluoroscopy for endovascular aortic repair. Int J Comput Assist Radiol Surg. 2018;13(8):1221-31.
12. Ronneberger O, Fischer F, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation2015 Mon, 13 Aug 2018 16:46:52 +0200; abs/1505.04597. Available from: http://arxiv.org/abs/1505.04597.