Cairo University

USING MID- AND HIGH-LEVEL VISUAL FEATURES FOR SURGICAL WORKFLOW DETECTION IN CHOLECYSTECTOMY PROCEDURES

By

Sherif Mohamed Hany Shehata

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE
in
Computer Engineering

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2016

USING MID- AND HIGH-LEVEL VISUAL FEATURES FOR SURGICAL WORKFLOW DETECTION IN CHOLECYSTECTOMY PROCEDURES

By

Sherif Mohamed Hany Shehata

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE
in
Computer Engineering

Under the Supervision of

Prof. Fathi Hassan Saleh
....
Professor of Computer Engineering
Computer Engineering Department
Faculty of Engineering, Cairo University

Dr. Nicolas Padoy
....
Assistant Professor
ICube laboratory
University of Strasbourg, France

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2016

USING MID- AND HIGH-LEVEL VISUAL FEATURES FOR SURGICAL WORKFLOW DETECTION IN CHOLECYSTECTOMY PROCEDURES

By

Sherif Mohamed Hany Shehata

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE
in
Computer Engineering

Approved by the Examining Committee:

Prof. Fathi Hassan Saleh, Thesis Main Advisor

Prof. Magda Bahaa Eldin Fayek, Internal Examiner

Prof. Samia Abdel Razek Mashaly, External Examiner
Professor at the Electronics Research Institute

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2016

Engineer: Sherif Mohamed Hany Shehata
Date of Birth: 14 / 09 / 1990
Nationality: Egyptian
E-mail: Sherif_mohamedhany@yahoo.com
Phone: +201112686899
Address 2 Montaser Bldgs, Haram st, Giza

Registration Date: 01 / 10 / 2014
Awarding Date: / / 2016
Degree: Master of Science
Department: Computer Engineering

Supervisors: Prof. Dr. Fathi Hassan Saleh
Dr. Nicolas Padoy Assistant Professor at ICube laboratory, University of Strasbourg, France

Examiners: Prof. Dr. Fathi Hassan Saleh (Thesis main advisor)
Prof. Dr. Magda Bahaa Eldin Fayek (Internal examiner)
Prof. Dr. Samia Abdel Razek Mashaly (External examiner)
, Professor at the Electronics Research Institute

Title of Thesis:

Using mid- and high-level visual features for surgical workflow detection in cholecystectomy procedures.

Key Words:

Cholecystectomy; Surgical workflow; Deformable part models; Convolutional neural network; Surgical tool detection

Summary:

We present a method that uses visual information in a Cholecystectomy procedure's video to detect the surgical workflow. While most related work relies on rich external information, we rely only on the endoscopic video used in the surgery. We fine-tune a convolutional neural network and use it to get mid-level features representing the surgical phases. Additionally, we train DPM object detectors to detect the used surgical tools, and utilize this information to provide discriminative high-level features. We present a pipeline that employs the mid- and high- level features by using one-vs-all SVMs followed by an HHMM to infer the surgical workflow. We present detailed experiments on a relatively large dataset containing 80 Cholecystectomy videos. Our best approach achieves 90% detection accuracy in offline mode using only visual information.

Acknowledgements

I would like to thank my main supervisor, Prof. Fathi Saleh, for his support and help throughout my masters. Without his support, I would not be able to reach this current stage.

A major part of this research was done while I was an intern at ICube laboratory in the University of Strasbourg in France. I worked in CAMMA group under supervision of Dr Nicolas Padoy. I would like to thank Dr Nicolas Padoy for his guidance throughout the experiment and for his feedback on my work afterwards.

I express my sincere gratitude to Andru P. Twinanda, PhD student in CAMMA group, for his contributions which helped me in reaching the current status in my research. First, he helped me in utilizing the features he used in one of his papers, which I use as a baseline for my work. Second, he extended the cholecystectomy dataset from 45 videos to 80 videos. Finally, he was involved in the evaluation process of the method proposed in this thesis.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AP	Average Precision
AUC	Area Under Curve
BOVW	Bag of Visual Words
CCA	Canonical Correlation Analysis
CNN	Convolutional Neural Network
CRF	Conditional Random Field
DPM	Deformable Parts Models
DTW	Dynamic Time Warping
HHMM	Hierarchical Hidden Markov Model
HMM	Hidden Markov Model
HOG	Histograms of Oriented Gradients
ILSVRC	ImageNet Large Scale Visual Recognition Challenge
LSTM	Long Short Term Memory
LSVM	Latent SVM
PCA	Principal Components Analysis
ReLU	Rectified Linear Unit
RFID	Radio-frequency identification technology
SGD	Stochastic gradient descent
SVM	Support Vector Machines

Abstract

We present a method that uses visual information from the video of laparoscopic cholecystectomy procedure to detect the surgical workflow. This task aims at recognizing the corresponding surgical phase for each frame of the laparoscopic video. In our method, we fine-tune a Convolutional Neural Network (CNN) and use it to extract mid-level features representing the surgical phases. Additionally, we train object detectors based on Deformable Parts Models (DPM) to detect the used surgical tools, then we utilize this information to provide discriminative high-level features. We present a pipeline that employs these mid- and high- level features to infer the surgical workflow. Our method uses one-vs-all Support Vector Machines (SVM) trained on the mid-level features to do initial assignment of phases' probabilities to each video frame. Afterwards, we concatenate the inferred phases' probabilities with the high-level features and feed these signals as observations for a Hierarchical Hidden Markov Model (HHMM). We use the HHMM to enforce the temporal constraints of phases' order and reach final recognition results.

Our major contribution is the set of visual features we use in our method. Most related work relies on rich external information regarding surgical tools usage. This information is generated using manual labeling or captured using additional equipment that are not available in common laparoscopic cholecystectomy procedures. On the contrary, our method relies only on visual features extracted from the laparoscopic video that is a basic component of all laparoscopic cholecystectomy procedures. The second contribution of our work comprises using a deep CNN in the task of detecting the surgical workflow. As far as we know, this is the first time that deep learning is used in this task. Using a deep CNN provides rich representations of the visual information inherent in the laparoscopic video, which helps in achieving state-of-the-art detection accuracy without relying on rich external information.

Furthermore, we present detailed experiments on a relatively large dataset, called Cholec80 dataset, which contains 80 laparoscopic cholecystectomy videos recorded and labeled at Strasbourg University. This dataset is 4-folds larger than the datasets used in previous studies. Our best approach, using only visual information, reaches state-of-the-art results on the Cholec80 dataset. Our approach achieves 90% detection accuracy in offline mode, where we process the full surgery video to infer the surgical workflow. As for the case of online mode, where video frames are processed without knowledge of future frames, our approach reaches 80% detection accuracy.

Chapter 1: Introduction

In recent years, the amount of technology used in medical applications increased dramatically. The goal of having fully automated surgeries have induced research in many directions. This thesis focuses on automatically detecting the surgical workflow in laparoscopic cholecystectomy surgical procedures. This would benefit in surgery automation, surgical skills assessment, and surgery summarization. In this chapter, we first introduce the laparoscopic cholecystectomy procedure and how it is performed. Next, we discuss the problem we are focusing on, surgical workflow detection, and explain our motivation and intended outcome. Then, we define our contribution in this thesis. Finally, we present the organization of this thesis.

1.1 Laparoscopic cholecystectomy surgical procedure

Cholecystectomy is the surgical removal of the gallbladder from the patient body. Laparoscopic cholecystectomy is the type of cholecystectomy in which the surgeons use small incisions to remove the gallbladder. Throughout laparoscopic cholecystectomy, a fiber optic camera is used to allow the surgeons to see inside the patient's abdomen through a small incision (figure 1.1). Cholecystectomy could be done using an open surgery, but the standard approach used in most cases is the laparoscopic cholecystectomy [1, 2, 3]. As any surgical operation, laparoscopic cholecystectomy may result in surgical complications. These complications include bile leak, bleeding, and bile duct injuries [4, 5, 6]. In some cases, surgeons convert the laparoscopic cholecystectomy to an open cholecystectomy to be able to handle the complications.

The surgery starts with preparations; first, the abdominal cavity is inflated using CO2. Inflation provides sufficient space for surgical operation, and provides visual clarity for the surgeons. Second, surgeons do four small incisions in patient's abdomen, and then insert a hollow tube, called trocar, through each incision. The trocars are surgeons' only access to the internal body. One of the trocars is the optical trocar, which is used to insert the laparoscopic camera. The other trocars are the operating trocars, which are used to insert surgical tools into

Figure 1.1: Screenshot from a video taken during an laparoscopic cholecystectomy procedure. It shows surgeons observing the laparoscopic video, which they utilize to see inside the patient's abdomen.

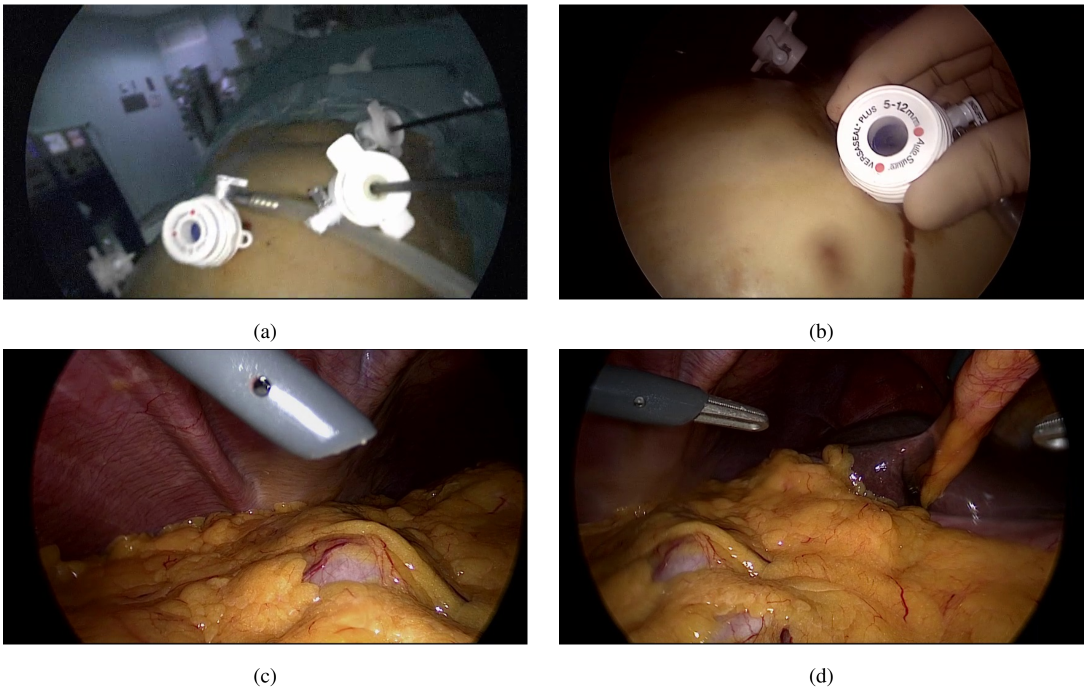
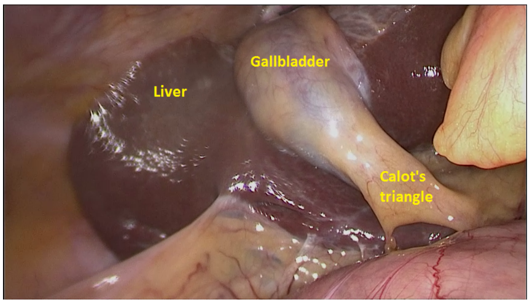


Figure 1.2: Screenshots from a cholecystectomy procedure showing trocars inside and outside the abdomen. Subfigure (a) shows the four trocars from outside the abdomen, with surgical tools inserted in the two trocars on the right. Subfigure (b) shows a close-up on one of the trocars outside the abdomen. Subfigures (c) and (d) show a trocar (the grey tube) inside the abdomen, with a tool inside it shown in (d).

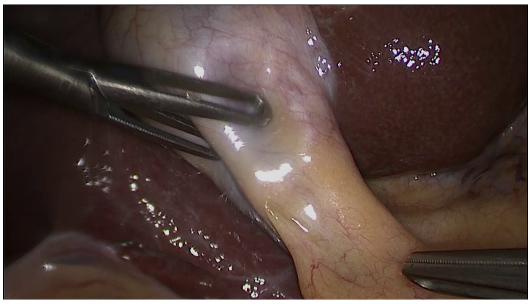
the abdomen. The main trocar is the one that contains tools that the surgeons use with their dominant hand. Figure 1.2 shows trocars inside and outside the abdomen in a laparoscopic cholecystectomy procedure. After inserting the four trocars, the main surgical steps start.

The gallbladder resides on the external surface of the liver. It is connected to the liver by the cystic duct and the cystic artery, which are located in the region called Calot's triangle (figure 1.3). After the preparations, the surgeon starts removing the fat from Calot's triangle. This clears the way for the surgeon to operate on the cystic duct and the cystic artery. Additionally, the surgeon cuts the tissues between cystic duct and the cystic artery to clear enough space for the tools used in next steps. After clearing the area, the surgeon uses a clipping tool to close the cystic artery and the cystic duct by applying multiple clips on them. Then the surgeon uses scissors to cut the cystic artery and the cystic duct. Clips are used to make sure that after the cutting step, bile will not leak from the cystic duct, and blood will not leak from the cystic artery. Since now the connections between the gallbladder and the liver are cut, the surgeon starts to detach the gallbladder. The surgeon cuts the tissues attaching the gallbladder to the liver bed until the gallbladder becomes fully detached. Finally, the gallbladder is put in a specimen bag, which is retracted through one of the trocars. The main surgical work is done and the surgical team works on closing incisions, and finalizing the surgery.

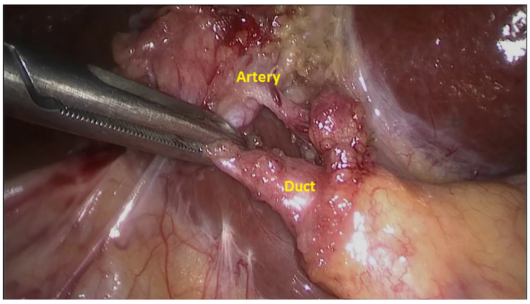
Throughout the surgery, the laparoscopic camera could get stains of blood or get blurred due to vapor condensation. In these cases, the surgeons retract the laparoscopic camera



(a)



(b)



(c)

Figure 1.3: Screenshots from cholecystectomy surgical procedure. They show (a) The gallbladder attached to liver bed, with Calot's triangle appearing below the gallbladder, (b) Calot's triangle before dissection, (c) Calot's triangle after dissection, showing the cystic duct (bottom) and the cystic artery (top). All screenshots are from the same surgical procedure.

outside the body and clean it. As a result, some parts of the cholecystectomy video does not show the abdominal cavity.

1.2 Surgical workflow detection

Having an intelligent system that detects performed phases of a surgical procedure has many benefits. This task, called surgical workflow detection, could help in monitoring the surgery's progress and its events. To be used in surgery monitoring, detection of surgical workflow needs to be performed online; the intelligent system needs to recognize current surgical phase while the surgery is being operated. Each surgical phase has its characteristics; as a result, each phase has different complication risks. A system for online surgical workflow detection could identify problems and risks in the operated surgery by identifying the current phase, then predicting the risks of complication specific to that phase. Furthermore, this online system could assist in setting the operating room schedule. Through monitoring the ongoing surgery, it could estimate the remaining time and notify the room management to adjust the schedule accordingly.

A system for surgical workflow detection has another set of applications if it works offline, where it processes the surgery's whole video after the surgery is completed. An offline system could be used for generating documentation of the surgery by identifying the operated