Departamento de
Informática e Ingeniería
de Sistemas
Universidad Zaragoza
1542

Trabajo Fin de Máster

Reconocimiento visual de imágenes de endoscopia con Deep Learning

Visual recognition with Deep Learning for endoscopic data

Clara Tomasini

Directora: Ana Cristina Murillo Arnal

Co-director: Luis Miguel Riazuelo Latas

Máster en Ingeniería biomédica
Departamento de Informática e Ingeniería de Sistemas
Escuela de Ingeniería y Arquitectura
Universidad de Zaragoza

22 de Junio de 2021

Índice general

Index II

1. Introducción **1**
1.1. Motivación 1
1.2. Contexto 2
1.3. Tareas y Objetivos 3
1.4. Estructura de la memoria 4

2. Aprendizaje automático en endoscopia **5**
2.1. Segmentación de herramientas con modelos supervisados 5
2.2. Análisis automático de vídeos de endoscopia con métodos no supervisados 7

3. Modelos supervisados para segmentación de herramientas **8**
3.1. Arquitecturas de red utilizadas 8
3.1.1. U-Net 8
3.1.2. TernausNet-11 9
3.1.3. LinkNet34 10
3.1.4. MiniNet 10
3.2. Sistema de segmentación implementado 11

4. Técnicas no supervisadas de análisis automático de datos **13**
4.1. Metodología de trabajo 13
4.2. Implementación 14
4.2.1. Extracción de características 14
4.2.2. Reducción de dimensionalidad 15
4.2.3. Visualización de los datos 15

5. Experimentación **17**
5.1. Entorno de experimentación 17
5.1.1. Datos utilizados 17
5.1.2. Entorno de trabajo 18
5.1.3. Métricas de evaluación 19
5.2. Segmentación de herramientas en vídeos de endoscopia 19
5.2.1. Re-entrenamiento de modelos en datos del proyecto . . . 19
5.2.2. Comparación de distintos modelos 23
5.3. Análisis no supervisado del contenido de vídeos de endoscopia . . 25
5.3.1. Visualización de los frames agrupados por tipo de herramienta 26

5.3.2. Visualización de los frames informativos y no informativos 28

6. Conclusión **30**
6.1. Conclusiones del Trabajo 30
6.2. Principales retos prácticos encontrados 30
6.3. Trabajo Futuro 31

Anexos **31**

A. Dataset EM **32**

B. Estructura del código **38**

Bibliografía **40**

Apéndice A

Dataset EM

En este apéndice se presentan frames elegidos aleatoriamente de cada vídeo utilizado del dataset EM. En el vídeo C capturado en el Hospital Clínico Universitario Lozano Blesa de Zaragoza se puede comprobar que aparece una herramienta que no aparece en ninguno de los otros vídeos, mientras que en los otros vídeos capturados en hospitales asociados al proyecto EndoMapper aparecen herramientas parecidas.

Vídeo A

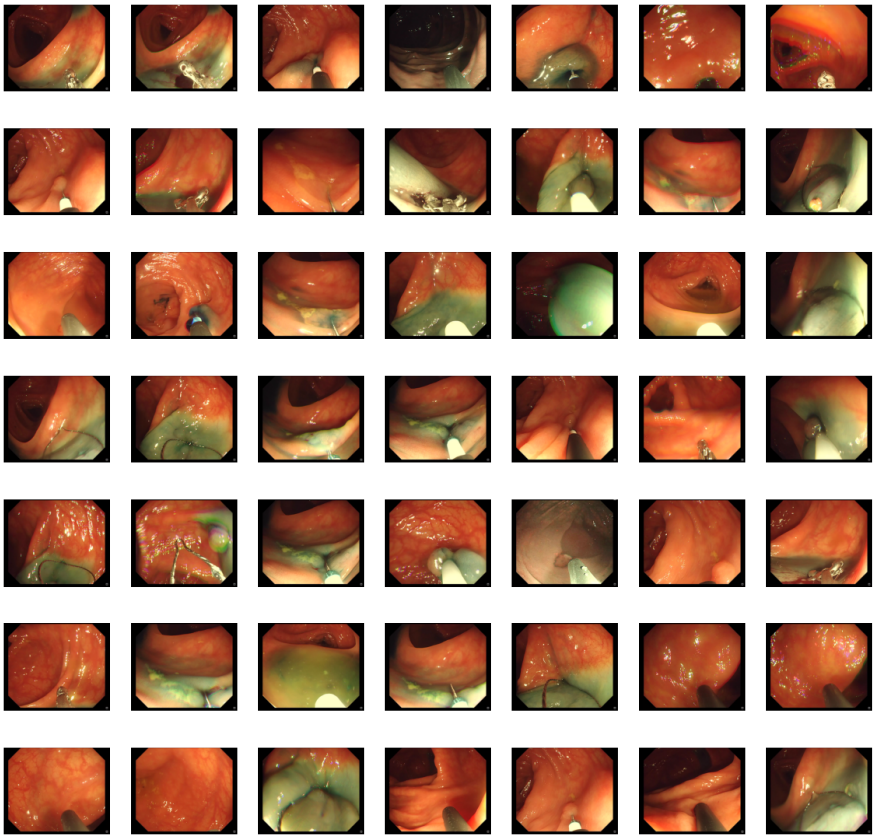

Figura A.1: 49 frames aleatorios del vídeo A

Vídeo B

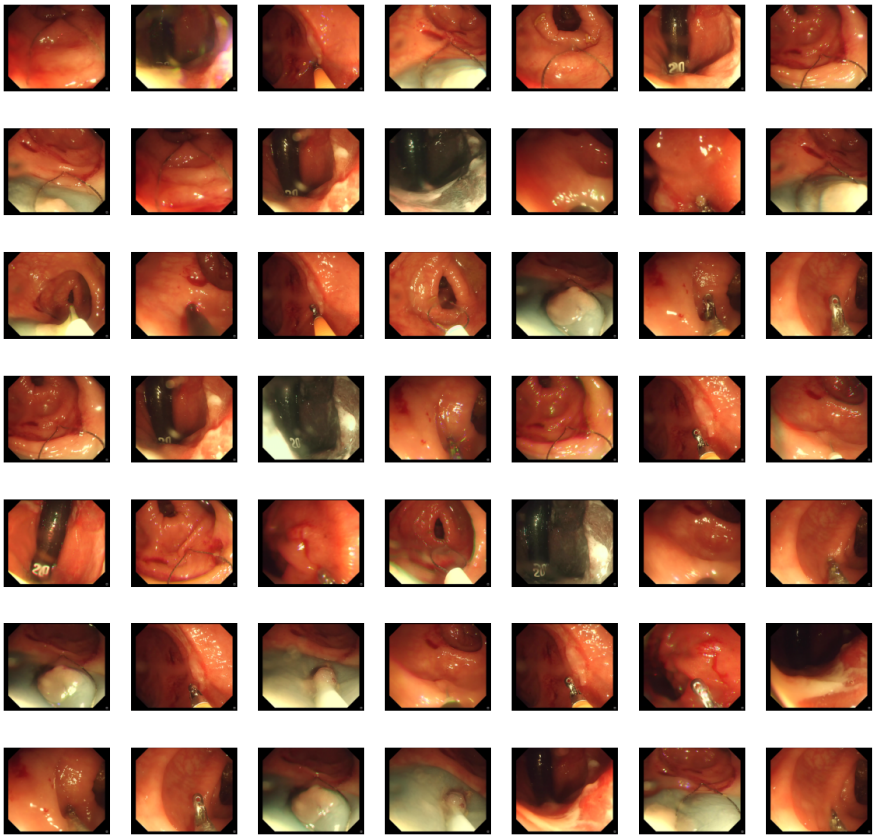

Figura A.2: 49 frames aleatorios del vídeo B

Vídeo C

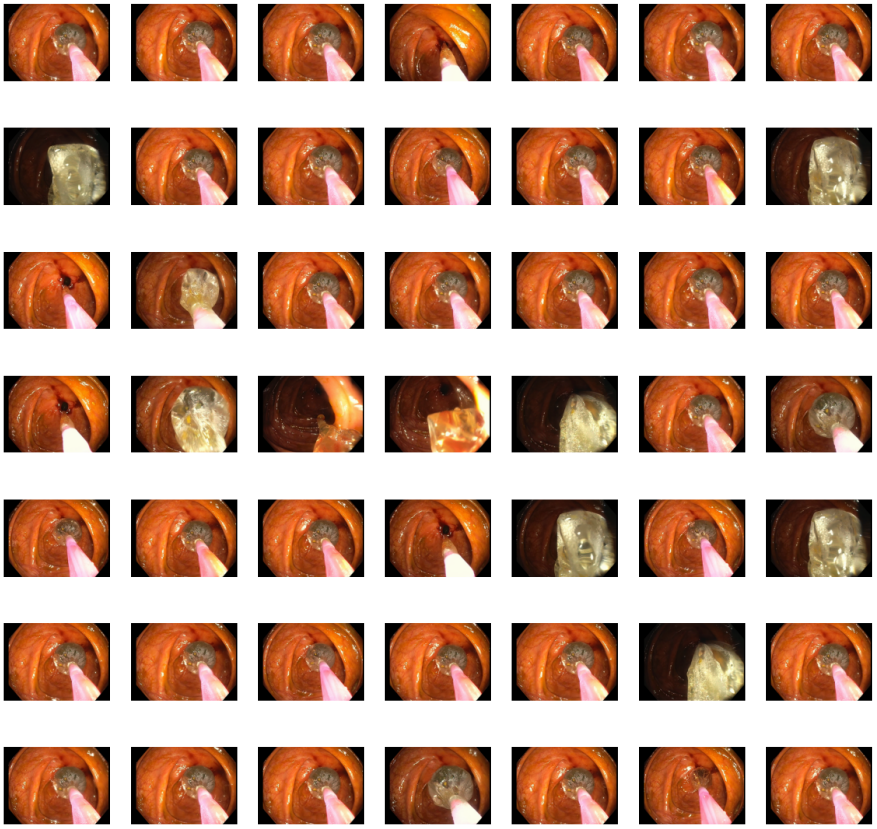

Figura A.3: 49 frames aleatorios del vídeo C

Vídeo D

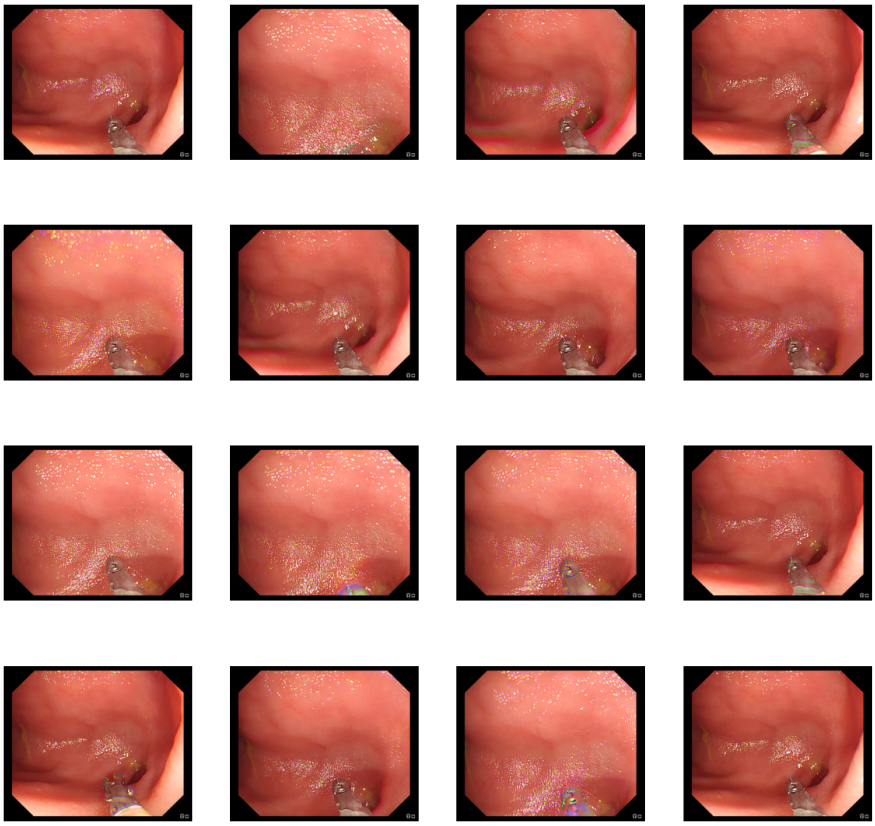

Figura A.4: 16 frames aleatorios del vídeo D

Vídeo E

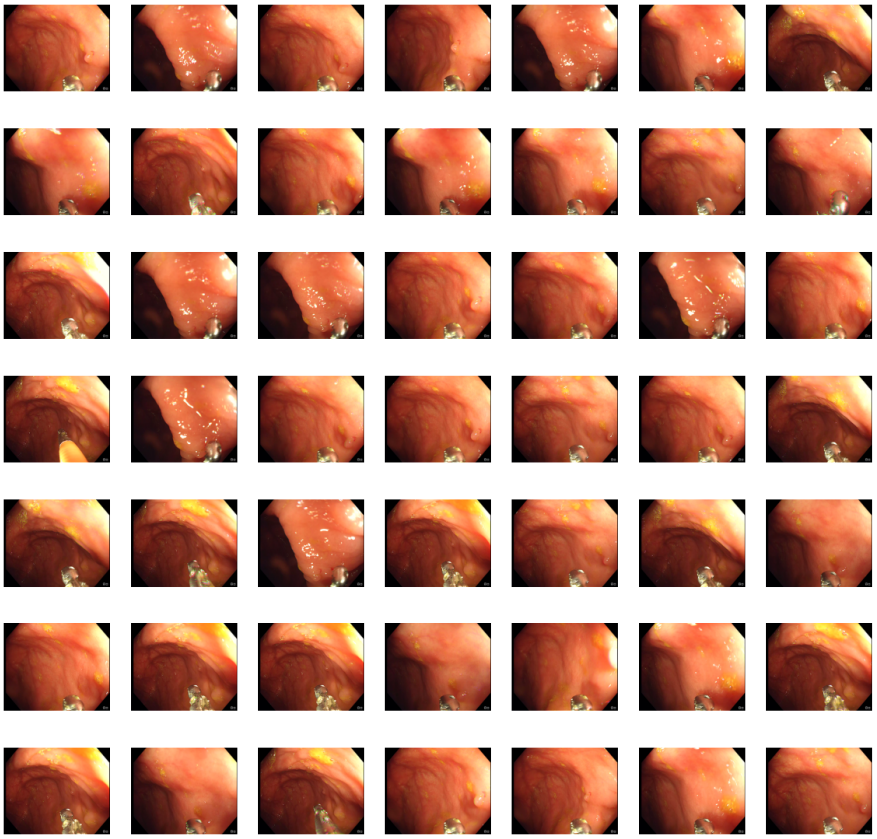

Figura A.5: 49 frames aleatorios del vídeo E

Apéndice B

Estructura del código

En este apéndice se presenta la estructura del código desarrollado para la primera parte de segmentación de este trabajo. El código está recopilado en un repositorio GitHub propio del proyecto, privado de momento. En este repositorio se encuentra el código de entrenamiento para los 4 modelos de segmentación, dividido entre 2 repositorios para MiniNet y los otros 3 modelos (UNet, TernausNet-11 y LinkNet-34, repositorio endovis_challenge). Se añade a continuación el fichero *Readme* del repositorio, que resume el trabajo hecho para la parte de segmentación y el código disponible.

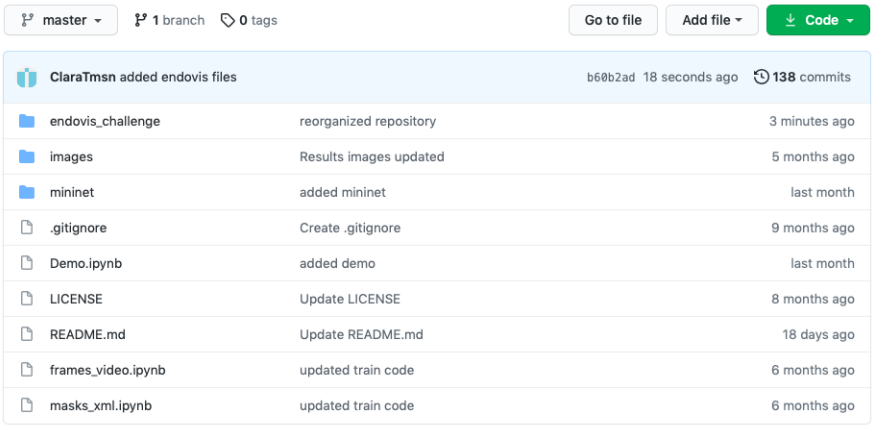

Figura B.1: Estructura del repositorio

README.md

toolSegmentation

This repository is built on a fork of project **robot-surgery-segmentation** (at https://github.com/ternaus/robot-surgery-segmentation), the official implementation of the paper

[1] *Automatic Instrument Segmentation in Robot-Assisted Surgery using Deep Learning*. Shvets, Alexey A., et al. IEEE Int. Conf. on Machine Learning and Applications. 2018.

The main goal is to obtain a tool segmentation model adapted to the requirements of the project **EndoMapper**.

Authors

Clara Tomasini, León Barbed, Ana Murillo, Luis Riazuelo, Pablo Azagra

How to run

Fine-tuned models are available at https://drive.google.com/drive/folders/1VOtD3U9jF4jPsPIZfU0LD3yXx0oI8d8-?usp=sharing. Best models obtained are Linknet models number 21 and 22 and Mininet.

XML files containing masks coordinates for UCL frames are available at
https://drive.google.com/drive/folders/1blv1w320DhIB3AW0UWRPtRgjlG4ephs5?usp=sharing

File *Demo.ipynb* provides an example of how to use the model in order to get a prediction for a given image using one of the models.

Results

All models were available pretrained on images similar to those of the Hamlyn dataset, and were then fine-tuned on more specific images from a different dataset (UCL). File *training.ipynb* shows how to fine-tune the models.

The following table shows several representative examples of the segmentations obtained for images both from the Hamlyn dataset [2] and from the project sequences (UCL and HCULB). The results use different models (UNet, TernausNet-11 and LinkNet-34) with the original and our fine-tuned versions. They show how the models fine-tuned with a few project labeled frames (just from one labeled sequence) adapt adequately to situations of our target domain (UCL and HCULB).

[2] *Three-dimensional tissue deformation recovery and tracking*. P. Mountney, D. Stoyanov, and G.-Z. Yang. IEEE Signal Processing Magazine, 27(4):14–24, 2010.

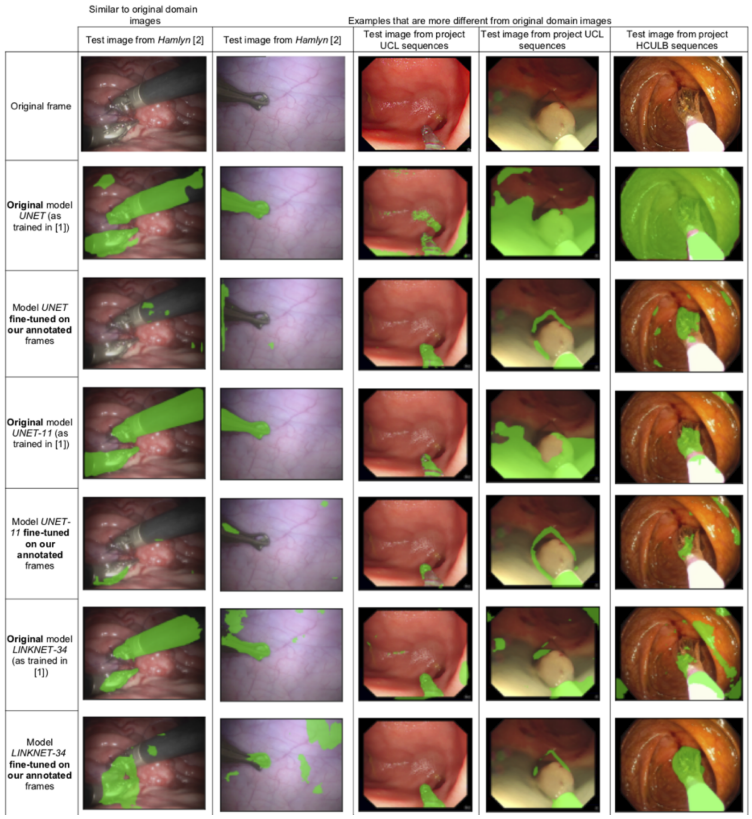

Figura B.2: Fichero *Readme* del repositorio

Bibliografía

[1] Alexander Rakhlin, Alexey Shvets, Vladimir Iglovikov, and Alexandr A. Kalinin. Deep convolutional neural networks for breast cancer histology image analysis. *Image Analysis and Recognition*, page 737–744, 2018. 1

[2] Aleksei Tiulpin, Jérôme Thevenot, Esa Rahtu, Petri Lehenkari, and Simo Saarakkala. Automatic knee osteoarthritis diagnosis from plain radiographs: a deep learning-based approach. *Scientific reports*, 8(1):1–10, 2018. 1

[3] Aji Resindra Widya, Yusuke Monno, Kosuke Imahori, Masatoshi Okutomi, Sho Suzuki, Takuji Gotoda, and Kenji Miki. 3d reconstruction of whole stomach from endoscope video using structure-from-motion. *arXiv e-prints*, pages arXiv–1905, 2019. 1

[4] EndoMapper. `https://sites.google.com/unizar.es/endomapper`, Abril 2021. 3

[5] Joonmyeong Choi, Keewon Shin, Jinhoon Jung, Hyun-Jin Bae, Do Hoon Kim, Jeong-Sik Byeon, and Namku Kim. Convolutional neural network technology in endoscopic imaging: artificial intelligence for endoscopy. *Clinical endoscopy*, 53(2):117, 2020. 5

[6] Hirotoshi Takiyama, Tsuyoshi Ozawa, Soichiro Ishihara, Mitsuhiro Fujishiro, Satoki Shichijo, Shuhei Nomura, Motoi Miura, and Tomohiro Tada. Automatic anatomical classification of esophagogastroduodenoscopy images using deep convolutional neural networks. *Scientific reports*, 8(1):1–8, 2018. 5

[7] Toshiaki Hirasawa, Kazuharu Aoyama, Tetsuya Tanimoto, Soichiro Ishihara, Satoki Shichijo, Tsuyoshi Ozawa, Tatsuya Ohnishi, Mitsuhiro Fujishiro, Keigo Matsuo, Junko Fujisaki, et al. Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images. *Gastric Cancer*, 21(4):653–660, 2018. 5

[8] Ruikai Zhang, Yali Zheng, Tony Wing Chung Mak, Ruoxi Yu, Sunny H Wong, James YW Lau, and Carmen CY Poon. Automatic detection and classification of colorectal polyps by transferring low-level cnn features from nonmedical domain. *IEEE journal of biomedical and health informatics*, 21(1):41–47, 2016. 5

[9] Bernd Münzer, Klaus Schoeffmann, and Laszlo Böszörmenyi. Content-based processing and analysis of endoscopic images and videos: A survey. *Multimedia Tools and Applications*, 77(1):1323–1362, 2018. 5

[10] Max Allan, Alex Shvets, Thomas Kurmann, Zichen Zhang, Rahul Duggal, Yun-Hsuan Su, Nicola Rieke, Iro Laina, Niveditha Kalavakonda, Sebastian Bodenstedt, et al. 2017 robotic instrument segmentation challenge. *arXiv preprint arXiv:1902.06426*, 2019. 6, 8, 17

[11] Alexey A Shvets, Alexander Rakhlin, Alexandr A Kalinin, and Vladimir I Iglovikov. Automatic instrument segmentation in robot-assisted surgery using deep learning. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 624–628. IEEE, 2018. 6, 8, 10, 11, 17, 20, 23

[12] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015. 6, 8, 9

[13] Vladimir Iglovikov and Alexey Shvets. Ternausnet: U-net with vgg11 encoder pre-trained on imagenet for image segmentation. *arXiv preprint arXiv:1801.05746*, 2018. 6, 9

[14] Abhishek Chaurasia and Eugenio Culurciello. Linknet: Exploiting encoder representations for efficient semantic segmentation. In *2017 IEEE Visual Communications and Image Processing (VCIP)*, pages 1–4. IEEE, 2017. 6, 10

[15] Alexandre Briot, Prashanth Viswanath, and Senthil Yogamani. Analysis of efficient cnn design techniques for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 663–672, 2018. 6

[16] Iñigo Alonso, Luis Riazuelo, and Ana C Murillo. Mininet: An efficient semantic segmentation convnet for real-time robotic applications. *IEEE Transactions on Robotics*, 36(4):1340–1347, 2020. 6, 10

[17] Wenju Du, Nini Rao, Dingyun Liu, Hongxiu Jiang, Chengsi Luo, Zhengwen Li, Tao Gan, and Bing Zeng. Review on the applications of deep learning in the analysis of gastrointestinal endoscopy images. *Ieee Access*, 7:142053–142069, 2019. 7

[18] Ryoma Bise, Kentaro Abe, Hideaki Hayashi, Kiyohito Tanaka, and Seiichi Uchida. Efficient soft-constrained clustering for group-based labeling. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 421–430. Springer, 2019. 7

[19] Jamil Ahmad, Khan Muhammad, Mi Young Lee, and Sung Wook Baik. Endoscopic image classification and retrieval using clustered convolutional features. *Journal of medical systems*, 41(12):1–12, 2017. 7

[20] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008. 7, 15

[21] Joris Guérin, Olivier Gibaru, Stéphane Thiery, and Eric Nyiri. Cnn features are also great at unsupervised classification. *arXiv preprint arXiv:1707.01700*, 2017. 14

[22] Carlos Oscar Sánchez Sorzano, Javier Vargas, and A Pascual Montano. A survey of dimensionality reduction techniques. *arXiv preprint arXiv:1403.2877*, 2014. 15

[23] EndoVis Challenge. Endovis'15 instrument subchallenge dataset, `https://opencas.webarchiv.kit.edu/?q=node/30`, 2015. 17

[24] Odin Vision. `https://odin-vision.com`, 2021. 18

[25] Google Colaboratory. `https://colab.research.google.com/notebooks/intro.ipynb`, 2021. 18

[26] MiniNet-v2. `https://github.com/Shathe/MiniNet-v2`, 2021. 23