# King's Research Portal

*Document Version*
Peer reviewed version

[Link to publication record in King's Research Portal](Link to publication record in King's Research Portal)

# Artificial intelligence for the real time classification of intrapapillary capillary loop patterns in the endoscopic diagnosis of early oesophageal squamous cell carcinoma: a proof of concept study.

**M Everson[1,2] LCGP Herrera[3], W Li[3], I Muntion Luengo[3], O Ahmad[1,3], M Banks [1,2], C Magee[1,2], D Alzoubaidi[1,2], HM Hsu[5], D Graham[1,2], T Vercauteren[3], L Lovat [1,2,3], S Ourselin[3], S Kashin[4], Hsiu-Po Wang[5], *Wen-Lun Wang[6], *RJ Haidry [1,2]**

[1]Division of Surgery & Interventional Science, University College London; [2]Department of Gastroenterology, University College Hospital NHS Foundation Trust, London; [3] Wellcome/EPSRC Centre for Interventional and Surgical Sciences (WEISS), UCL, London, UK; [4]Yaroslavl Regional Cancer Hospital, Yaroslavl, Russia; [5]Department of Internal Medicine, National Taiwan University, Taipei, Taiwan; [6] Department of Internal Medicine, E-Da Hospital/I-Shou University, Kaohsiung, Taiwan

**Conflicts of interest:** RJH has received research grant support from Pentax Medical, Cook Endoscopy, Fractyl Ltd, Beamline Ltd and Covidien plc to support research infrastructure. No other conflicts of interest

**All correspondence to:** *Joint senior authors*

Dr Rehan Haidry, Consultant Gastroenterologist, Director of Endoscopy, 235 Euston Road, London, United Kingdom, NW1 2BU
*Email: r.haidry@ucl.ac.uk*

Wen-Lun Wang, MD, PhD. Department of Internal Medicine, E-Da Hospital/I-Shou University, Kaohsiung, Taiwan No. 1, Yida Road, E-mail:warrengodr@gmail.com

**ME, LCGPH, HPW, WLW, RJH** - study concept and design; acquisition of data; analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for important intellectual content; statistical analysis; obtained funding; administrative, technical, or material support; study supervision

**WL, IML, TV, SO** - analysis and interpretation of data; statistical analysis

**OA, MB, CM, DA, HMH, DG, LL, SO, SK** - critical revision of the manuscript for important intellectual content; material support

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

## Abstract

### Background

Intrapapillary capillary loops (IPCLs) represent an endoscopically visible feature of early squamous cell neoplasia (ESCN), which correlate with invasion depth – an important factor in the success of curative endoscopic therapy. IPCLs visualised on ME with narrow band imaging (ME-NBI) can be used to train convolutional neural networks (CNNs) to detect the presence and classify staging of ESCN lesions.

### Methods

7046 sequential HD ME-NBI images from 17 patients (10 ESCN, 7 normal) were used to train a CNN. IPCL patterns were classified by three expert endoscopists according to the Japanese Endoscopic Society (JES) classification. Normal IPCLs were defined as type A, abnormal as B1-3. Matched histology was obtained for all imaged areas.

### Results

This CNN differentiates abnormal from normal IPCL patterns with 93.7% accuracy (86.2-98.3%), sensitivity and specificity for classifying abnormal IPCL patterns of 89.3% (78.1-100%) and 98% (92-99.7%) respectively. Our CNN operates in real time with diagnostic prediction times between 26.17-37.48ms.

### Conclusion

Our novel and proof of concept application of computer-aided endoscopic diagnosis (CAED) shows that a CNN can accurately classify IPCL patterns as normal or abnormal. This system could be used as an in-vivo, real-time clinical decision support tool for endoscopists assessing and directing local therapy of ESCN.

**Established knowledge on the subject**

- The JES classification provides a clinically useful and simplified system by which to predict histology in early squamous lesions based on IPCL patterns
- Artificial intelligence may have a role in improving clinicians endoscopic recognition of IPCL patterns – no previous study has yet been reported in the literature

**Significant new findings in this study**

- We introduce a novel artificial intelligence (AI) system capable of classifying IPCL patterns as neoplastic or non-neoplastic in real-time, using ME-NBI images acquired during endoscopy.
- Our system demonstrates an impressive accuracy and could potentially assist clinicians in the recognition of neoplastic tissue, as well as providing a clinical decision support tool to guide endoscopists considering whether lesions are amenable to endoscopic resection.
- Our system provides clinically interpretable images of which IPCLs are abnormal to assist clinicians delineating lesions
- Our system provides a performance benchmark for the assessment of other AI systems designed for this purpose in the future.

## Introduction

Oesophageal squamous cell carcinoma is the eighth most common cause of cancer worldwide and the sixth most common cause of cancer deaths[1], with the highest incidence across a 'cancer belt' extending from East Africa, the Middle East to China and Japan[1–3]. Gastroscopy remains the investigation of choice for the diagnosis of early squamous cell neoplasia (ESCN). The endoscopic features of these early lesions are subtle and easily missed during endoscopy - studies show a significant miss rate for UGI cancers on endoscopies undertaken within three years of diagnosis[4]. Early detection and accurate characterisation of ESCN lesions is vital, in order to predict histology and guide intervention. Lesions confined to the mucosa have low rates of local lymph node metastasis (<2%) compared to lesions invading the submucosa (8-45.9%) and so are amenable to endoscopic therapy[5,6,7,8]. Endoscopic therapy, through endoscopic mucosal resection (EMR) or submucosal dissection (ESD) is associated with excellent rates of 5 year survival[9] and spares patients the attendant morbidity and mortality of oesophagectomy after a late diagnosis[10,11].

Intrapapillary capillary loops (IPCLs) are microvessels that were first characterised using magnification endoscopy[12,13]. They are now an established marker of ESCN and changes in their morphology correlate with the invasion depth of ESCN[14,15]. Normal IPCLs arise from the submucosal vessels, run adjacent to basal layer of the oesophageal epithelium and are seen as fine calibre looped structures on magnified endoscopy. As an ESCN progresses there is stepwise destruction of the oesophageal wall architecture, which manifests as morphologic changes in the ICPL pattern. Initially IPCLs become more tortuous and dilated. As the ESCN progresses and interrupts the mucosal layer, the IPCLs lose their looped structure and form linear dilated vessels. Avascular areas form between these non-looped vessels which also correspond to invasion depths deep to the mucosal layer. As the ESCN invades the deeper submucosal layer, the IPCLs are almost completely obliterated. In their place neovascularisation occurs with the formation of tortuous, grossly dilated and non-looped vessels[16–18].

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Advanced endoscopic imaging modalities, such as narrow band imaging (NBI) in combination with magnification endoscopy, afford improved visualisation of subtle microvascular patterns in the oesophageal mucosa of patients with ESCN[19]. Several classifications have been proposed to define the abnormal ICPL morphologies that correlate with histological invasion depth of ESCN lesions. Classifications described by Inoue et al[12,13]. and Arima et al.[20]; associate progressive morphological abnormalities in IPCLs with deeper invasion of the ESCN[14]. These classifications are complex and require a high degree of interpretation by endoscopists – as such their utility in a clinical setting is debatable.

The recently published Japanese Endoscopic Society (JES) IPCL classification is a comparably simplified system, allowing easy recognition of ESCN by endoscopists[21,15]. The JES classification has become widely accepted in areas of high prevalence such as China and Japan. Importantly each ICPL subgroup corresponds with high accuracy to a given histological grade and invasion depth of ESCN – with increasing irregularity of the IPCL patterns representing more advanced, invasive disease[15] (figure 1a and 1b). The accuracy of the JES classification as reported by Oyama et al is high compared to other classifications – with the overall accuracy for histology prediction 90.5% across type B1-3. Overall accuracies for histology prediction were 91.9%, 93.4% and 95.9% for type B1, B2 and B3 IPCL patterns respectively[15]. Kim et al also report excellent interobserver agreement using the JES classification[22].

Computer-aided endoscopic diagnosis (CAED) with AI could provide a useful adjunct to endoscopists assessing ESCN lesions; a system capable of recognising and highlighting the presence of ESCN could improve detection rates. CAED using convolutional neural networks (CNNs) has shown great potential in a range of medical specialties, but its use in the recognition of endoscopically detected neoplasia is in its infancy[23]. We propose that the distinct patterns observed in IPCL morphology could provide the input data to train a CNN to classify such microvasculature patterns as normal or abnormal. Such a system could improve and automate the characterisation and prediction of ESCN invasion depth, which would be particularly useful in settings where endoscopists are unfamiliar with the endoscopic appearance of early squamous lesions. Furthermore, a downstream application of such a validated system, trained to differentiate IPCL subtypes has a role as a support tool for endoscopists to triage lesions as either amenable or unamenable to EET – based on the predicted invasion depth deduced from the IPCL patterns, reducing procedure time and preventing endoscopic resection in patients for whom it would be either inappropriate or futile.

**Study aims:**

1. To develop a novel AI system that can classify IPCL patterns as normal or abnormal in **endoscopically resectable lesions** (<SM1 invasion) in real time on videos acquired during magnification endoscopy

2. To develop a methodology that can be used to develop further AI systems, for use in vivo, that accurately predicts the grade of a lesions histology and invasion depth based on IPCL patterns

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

## Methods

### *Patient recruitment, inclusion/exclusion criteria*

Patients attending two ESCN referral centres in Taiwan (National Taiwan University Hospital and E-Da Hospital) were recruited with consent. Patients were only included if pathological samples (EMR, ESD, oesophagectomy specimens) were acquired at the time of their endoscopy or subsequent surgery. Patients with oesophageal ulceration were excluded. Our study complied with the Declaration of Helsinki. The Institutional Review Board of E-Da Hospital approved this study (IRB number: EMRP-097-022. July 2017).

### *Endoscopic procedures and video acquisition*
Gastroscopies were performed by two expert endoscopists (WLW, HPW), both performing >50 ESCN assessments per year. Endoscopies were performed using a HD ME-NBI GIF-H260Z endoscope, with Olympus Lucera CV-290 processor (Olympus, Tokyo, Japan). The oesophageal mucosa was cleaned with a Simethicone solution prior to the interrogation of the IPCL patterns using magnification endoscopy with narrow band imaging (ME-NBI). Magnification endoscopy was performed on areas of interest at 80-100x magnification. After assessment of the lesion, pathological samples of the imaged area were acquired through EMR or ESD. Histopathological analysis was undertaken by two expert gastrointestinal pathologists and reported according to the Vienna classification system[24].

### *Labelling ME-NBI videos and establishing an expert consensus*
HD endoscopic videos were reviewed independently by three expert upper gastrointestinal endoscopists with extensive experience in the endoscopic assessment of ESCN lesions (WLW, RJH and HPW). The visualised IPCL patterns were classified by consensus based on the JES classification system. Classifications were then correlated with histology from the imaged area before a final classification was assigned to each image. Type A IPCLs were considered normal, type B1, 2 and 3 IPCLs were considered abnormal. Videos were sampled at 30fps to generate sequential still images which were stored as the lossless .png format.

*Rationale for selecting images as a data input*

An AI system for use in-vivo needs to be able to classify IPCL patterns at video rate. A video is simply a rapid sequence of images and can be broken down into individual frames. Individual frames provide more data to train the CNN and is acceptable as a data input for training AI networks for video classification.

*Image quality control*

Images were manually quality controlled by a senior clinician study member (ME); those that were blurred, or if there was obscuration of the mucosa by blood or mucus were removed from both the training and testing datasets. The black borders surrounding the endoscopic image were also removed

*Formation of image datasets*

The full dataset consisted of 7046 images (8193 excluded), with resolutions ranging from 458x308 to 696x308 pixels. We employed 5-fold cross validation to generate five distinct datasets with different combinations of images (figure 2a). 5-fold validation serves two purposes; first, to maximise the number of images that can be used for training, validation and testing. For any given fold, images of patients used within the training dataset were not used in the validation or testing dataset (summarised in table 1 and 2). Secondly each fold generates a distinct CNN, by developing multiple networks simultaneously, one can interrogate that the designed network architecture performs well across a range of patients and is not simply performing by 'fluke' on one particular combination of patients images.

*Convolutional neural network design*

A full description of our technical methodology has previously been described by our group[25] (figure 2b). Quality controlled ME-NBI images of ESCNs were used as the input data. Each image was labelled as representative of a specific IPCL subtype. In training, input images at various magnifications are passed through layers of the CNN. Images pass through 5 CNN layers, with a fixed learning rate of $1^{-6}$, momentum of 0.9 and a batch size of 1. Stochastic Gradient Descent was the optimiser. All folds were trained with a maximum iteration of 4x the number of images in the training set. Explicit class activation maps (eCAMs) are generated to depict visually areas of each image that the CNN finds most discriminative when classifying them as normal or abnormal. Images were not resized but global average pooling was employed after the decoder and prior to classification/ In a final output layer the CNN classifies the IPCL pattern as normal or abnormal based on a 1x1 convolution. If the prediction is correct compared to the clinician labelled input image, filters are conserved. Once training is complete, new, unseen images are inputted into the CNN and its classification performance is assessed.

*Statistical analysis*

Accuracy, F1 scores (a weighted average of precision and sensitivity), sensitivity and specificity for abnormal IPCL classification were calculated. Power calculations were not

deemed necessary for this study since it is entirely novel and a proof of concept work. We are attempting only to benchmark the currently achievable diagnostic performance of AI systems in the classification of IPCL patterns.

## Results

### *Patient characteristics*

17 patients were included in our dataset; 10 with ESCN and 7 with a normal squamous oesophagus. All patients were deemed to have endoscopically resectable lesions, a summary of the histological invasion depth given in table 3.

### *CNN performance for IPCL classification*

Our CNN operates at video rate, capable of classifying sequential HD-images acquired from endoscopic videos. The classification interval time varied according to the size of the images analysed, but ranged from 26.17ms to 37.48ms. The CNN demonstrated a mean accuracy for the differentiation of abnormal IPCL patterns (B1/B2/B3) from normal type A patterns of 93.3% (range 86.2-98.3%). The average F1 score (a weighted average of precision and sensitivity) for identifying ESCN based on the IPCL pattern was similarly high at 92.7% (range 85.4-98.2%). Our algorithm achieved a sensitivity and specificity for abnormal IPCL classification of 89.7% (range 78.1-100%) and 96.9% (range 92-99.7%) respectively. There was some variability in performance statistics between folds, which likely represents that the CNN is still not adept at distinguishing all of the varied IPCL features that represent abnormal tissue (Type B1/2/3). While our dataset is the largest in the published literature it is highly likely that more images are required to further train the network to recognise the full spectrum of variability in IPCL patterns. Our CNN performance statistics are summarised in table 4.

### *Analysis of explicit class activation maps (eCAMs) provides a clinically interpretable neural network output*

The distribution of abnormal IPCL patterns are heterogeneous between images, eCAMs were therefore generated to elucidate which visual features within the images the CNN was using to

discriminate healthy and unhealthy tissue. They also provide a clinically interpretable network output. This served as a further validation step to ensure that the CNN was using the IPCL patterns in order to classify images as normal or abnormal, rather than other subtle discriminative features. Analysis of the eCAMs patterns (figure 3), offers several interesting insights into the decision-making process of our CNN. Firstly, as expected, our CNN appears to use the IPCL patterns as its most discriminative feature. Secondly it is capable of recognising discrete areas of abnormal IPCL patterns within otherwise normal mucosa and ignores other uninformative features such as specular reflections. Interestingly, the CNN did not use avascular areas between the grossly abnormal B2 IPCLs seen in figure, but remained able to discriminate the abnormal vessels highly selectively. We also note that within some images of healthy tissue the CNN has a tendency to classify the deep submucosal vessels as abnormal – which may potentially lead to classifications of normal tissue as abnormal.

## Discussion

We introduce the first application of CAED and AI for the accurate, video rate characterisation of ESCN through the classification of intrapapillary capillary loops seen during magnification endoscopy based on the JES system.

Accurate assessment and characterisation of ESCN lesions is vital in order to predict histology and invasion depth to guide appropriate intervention. Lesions confined to the mucosa have low rates of local lymph node metastasis (<2%) and so are amenable to endoscopic therapy[5,6,7,8], with reported 5 year survivals of 75-100% for mucosal ESCN[9].

Several classification systems exist to assist clinicians with the characterisation of IPCL patterns. Inoue et al (2001) were the first to propose a five-part classification of IPCL patterns[26]. Type I-III IPCL patterns were associated with normal mucosa, inflammation or LGIN. Type IV IPCLs were suggestive of HGD. Type V IPCLs were subdivided into $V_1$ (M1 carcinoma in situ), $V_2$ (M2 carcinoma in situ), $V_3$ (M3 or early SM1 invasion) and $V_4$ (invasion to at least SM2) based on the progression of abnormal IPCL morphology[26]. In one study, the Inoue classification enabled identification of early mucosal lesions (M1-2) with 89.5% sensitivity[14]. The Inoue classification sensitivity for prediction of ESCN invasion into SM1 and SM2 was much lower at 58.7% and 55.8% respectively[14]. Arima et al (2005) proposed an alternative four-part classification system with type 1, 2, 3 and 4 IPCLs representing normal mucosa, inflammatory changes, M1-2 carcinoma in situ and >M3 invasion respectively[20].

We elected to use the JES IPCL classification of ESCN as the input and output for our CNN as it is the most contemporary and accurate system and is widely used in areas of high ESCN prevalence. We believe that a CNN capable of using the JES classification will remain clinically relevant as it becomes more widely used. We suggest that the classifications proposed by Inoue[26] and Arima[20] are complex and may not be as intuitive to use.

Our study, the first reported use of CNNs for this purpose, uses sequential still images captured from HD videos of endoscopic examinations to train a CNN to characterise ESCN lesions based on the IPCL patterns visualised at endoscopic assessment. This CAED platform achieved an accuracy of 93.7% for the characterisation of abnormal IPCL patterns. The sensitivity and specificity of our CNN compares favourably to other published work at 89.7% and 96.9% respectively. Our CNN operates at video rate, but it should be noted that more work is required for this to be put to clinical use; including training a CNN to remove uninformative frames without the need for manual selection.

We envisage that our CNN could assist endoscopists in the recognition of abnormal IPCL patterns in patients undergoing assessment for suspected ESCN and consideration of endoscopic treatment.  Such a system could shorten the time required for endoscopic evaluation, reduce interobserver variability and importantly inform in-vivo clinical decision making on what lesions require or are amenable to endoscopic resection. Wang et al (2017) demonstrate that in a non-expert panel of endoscopists accuracy of histology prediction using the JES classification ranged from 48-57% after a short training program[27]. Kim et al (2017) also demonstrated an overall accuracy of 78.6% for the correct prediction of histology[22].  Our CNNs overall accuracy of 93.7% therefore compares very favourably to this. We would therefore expect that our system would most benefit for non-expert endoscopists or centres where the endoscopic assessment of ESCN is less common and so clinicians may feel less confident in recognising and classifying IPCL patterns.

Our study has several limitations, all of which we argue are to be expected in a feasibility study, but have nonetheless attempted to mitigate. Firstly, our patient sample size is small, using 7046 images generated from 17 patients. However, this is currently the largest image dataset used to train a CNN for this specific application reported in the literature. To mitigate this, we utilised 5-fold cross validation, as outlined in table 4, in order to benchmark the ability of the CNN to recognise a range of IPCL patterns under different operating conditions. This method ensures that our network was trained and tested on all available images, whilst ensuring that no images were used for both training and testing. Our novel use of eCAMs also acts as an additional form of validation; we have demonstrated that the ICPL patterns are the visual features the network finds discriminative when classifying images. We therefore expect that with greater patient numbers we should continue to see such high accuracies for IPCL characterisation.

Secondly, our system accurately differentiates normal (type A) compared to abnormal (type B1-3) IPCL patterns, but is not yet able to differentiate individual subtypes with sufficient accuracy for clinical use. This is comparable to other studies using the Inoue and Arima classifications which reported less accurate characterisation of >SM1 lesions based on IPCL patterns. We have only included images of type A, B1 and B2 IPCLs as an input to our CNN for this study, since we currently seek to develop a system that highlights abnormal areas of potentially resectable lesions to clinicians in the first instance.

Lastly, our gold standard used three expert clinicians. We feel that for a pilot study this is sufficient to provide accurate ground truth to develop a CNN, particularly since the consensus

classifications were correlated with histological results. We recognise that future work is needed to validate this CNNs performance against a larger panel of clinicians of varying experience before it could be used in a clinical environment. It is important to note that published work on the JES classification reports interobserver variability of between 0.61[27] and 0.86[22]. A CNN developed to classify IPCL patterns should therefore only be expected to obtain this level of variability during external validation. An automated method of removing uninformative frames may also reduce the selection bias introduced by a clinician removing uninformative frames.

Further work currently underway aims to produce a CNN capable of more precise classification based on the individual JES subgroups. Such a validated system would also improve the characterisation of ESCN and could allow clinicians to triage, in real—time, lesions that would be amenable to EET, compared to lesions which would be too advanced to be cured endoscopically. This would prevent patients with advanced disease undergoing unnecessary endoscopic therapy, while ensuring that they are referred promptly for surgical management.

Word count 2997 – excluding abstracts, figures, references

## Acknowledgments

## References

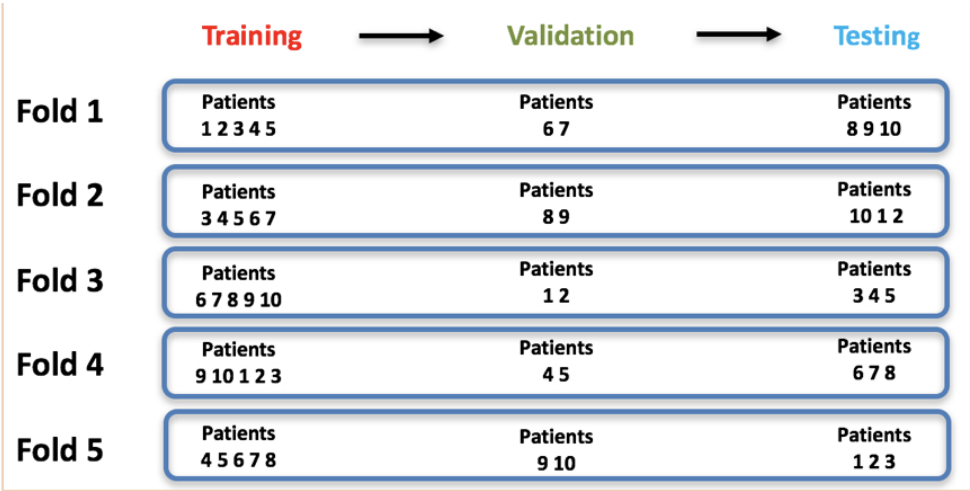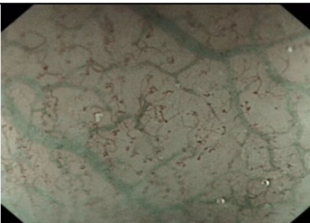1.    Zhang, Y. Epidemiology of esophageal cancer. *World J. Gastroenterol.* **19,** 5598–5606 (2013).
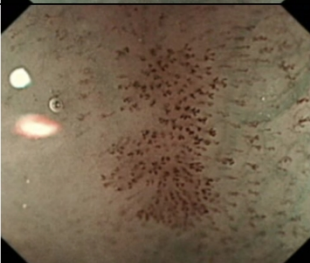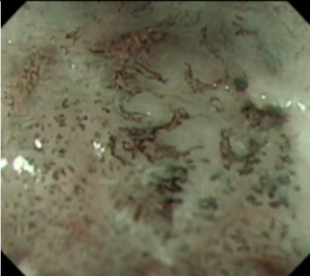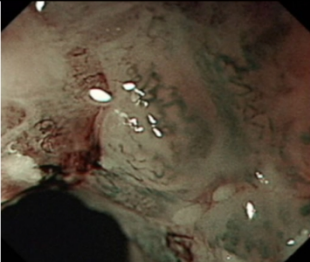
2.    Eslick, G. D. Epidemiology of Esophageal Cancer. *Gastroenterol. Clin. North*

*Am.* **38,** 17–25 (2009).

3.    Taylor, P. R., Abnet, C. C. & Dawsey, S. M. Squamous dysplasia – the precursor lesion for esophageal squamous cell carcinoma. *Cancer Epidemiol. Biomarkers Prev.* **22,** 540–552 (2013).

4.    Menon, S. & Trudgill, N. How commonly is upper gastrointestinal cancer missed at endoscopy? A meta-analysis. *Endosc. Int. Open* **2,** E46–E50 (2014).

5.    Cho, J. W. *et al.* Lymph Node Metastases in Esophageal Carcinoma: An Endoscopist's View. *Clin. Endosc.* **47,** 523–529 (2014).

6.    Shimada, H. *et al.* Impact of the Number and Extent of Positive Lymph Nodes in 200 Patients with Thoracic Esophageal Squamous Cell Carcinoma after Three-field Lymph Node Dissection. *World J. Surg.* **30,** 1441–1449 (2006).

7.    Sgourakis, G. *et al.* Detection of lymph node metastases in esophageal cancer. *Expert Rev. Anticancer Ther.* **11,** 601–612 (2011).

8.    Kodama, M. & Kakegawa, T. Treatment of superficial cancer of the esophagus: a summary of responses to a questionnaire on superficial cancer of the esophagus in Japan. *Surgery* **123,** 432–439 (1998).

9.    Shimizu, Y. *et al.* Long-term outcome after endoscopic mucosal resection in patients with esophageal squamous cell carcinoma invading the muscularis mucosae or deeper. *Gastrointest. Endosc.* **56,** 387–390 (2018).

10.   Inoue, H., Minami, H., Kaga, M., Sato, Y. & Kudo, S. Endoscopic Mucosal Resection and Endoscopic Submucosal Dissection for Esophageal Dysplasia and Carcinoma. *Gastrointest. Endosc. Clin. N. Am.* **20,** 25–34 (2010).

11.   Oyama, T. *et al.* Endoscopic Submucosal Dissection of Early Esophageal Cancer. *Clin. Gastroenterol. Hepatol.* **3,** S67–S70 (2018).

12.   INOUE, H. *et al.* Ultra-high magnification endoscopy of the normal esophageal mucosa. *Dig. Endosc.* **8,** 134–138 (1996).

13.   INOUE, H. *et al.* Ultra-high magnification endoscopic observation of carcinoma

in situ of the esophagus. *Dig. Endosc.* **9,** 16–18 (1997).

14. Sato, H. *et al.* Utility of intrapapillary capillary loops seen on magnifying narrow-band imaging in estimating invasive depth of esophageal squamous cell carcinoma. *Endoscopy* **47,** 122–128 (2015).

15. Oyama, T. *et al.* Prediction of the invasion depth of superficial squamous cell carcinoma based on microvessel morphology: magnifying endoscopic classification of the Japan Esophageal Society. *Esophagus* **14,** 105–112 (2017).

16. Kaga, M., Inoue, H., Kudo, S.-E. & Hamatani, S. Microvascular architecture of early esophageal neoplasia. *Oncol. Rep.* **26,** 1063–1067 (2011).

17. Inoue, H. *et al.* Magnification endoscopy in esophageal squamous cell carcinoma: a review of the intrapapillary capillary loop classification. *Ann. Gastroenterol. Q. Publ. Hell. Soc. Gastroenterol.* **28,** 41–48 (2015).

18. Kumagai, Y., Toi, M., Kawada, K. & Kawano, T. Angiogenesis in superficial esophageal squamous cell carcinoma: magnifying endoscopic observation and molecular analysis. *Dig. Endosc.* **22,** 259–267 (2010).

19. Gono, K. *et al.* Appearance of enhanced tissue features in narrow-band endoscopic imaging. *J. Biomed. Opt.* **9,** 568–577 (2004).

20. Arima, M., Tada, M. & Arima, H. Evaluation of microvascular patterns of superficial esophageal cancers by magnifying endoscopy. *Esophagus* **2,** 191–197 (2005).

21. Oyama, T. A new classification of magnified endoscopy for superficial esophageal squamous cell carcinoma. *Esophagus* **8,** 247–251 (2011).

22. Kim, S. J. *et al. New magnifying endoscopic classification for superficial esophageal squamous cell carcinoma. World Journal of Gastroenterology* **23,** 4416–4421 (Baishideng Publishing Group Inc, 2017).

23. Zhang, C. *et al.* Tu1217 The Use of Convolutional Neural Artificial Intelligence

Network to Aid the Diagnosis and Classification of Early Esophageal Neoplasia. A Feasibility Study. *Gastrointest. Endosc.* **85,** AB587-AB588 (2018).

24.   Schlemper, R. J. *et al.* The Vienna classification of gastrointestinal epithelial neoplasia. *Gut* **47,** 251–255 (2000).

25.   Garcia-Peraza-Herrera, L. *et al.* Interpretable Fully Convolutional Classification of Intrapapillary Capillary Loops for Real-Time Detection of Early Squamous Neoplasia. *Arxiv* (2018). doi:arXiv:1805.00632

26.   Inoue, H. Magnification endoscopy in the esophagus and stomach. *Dig. Endosc.* **13,** (2001).

27.   Wang, W.-L. *et al.* A training program of a new simplified classification of magnified narrow band imaging for superficial esophageal squamous cell carcinoma. *J. Gastroenterol. Hepatol.* n/a-n/a doi:10.1111/jgh.14071

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

| JES type | | Typical morphology | Typical histology |
|---|---|---|---|
| A |  | Small, non-dilated, fine calibre looped vessels with no gross abnormality. Some vessels may take a more elongated form with inflammation or LGIN. Submucosal vessels may be visible with the background mucosa of a uniform colour | Normal/LGIN |
| B1 |  | Subtle abnormalities in IPCLs observed. Increased tortuosity, increased vessel calibre and density but looped structure is retained. There may be a brown hue to the mucosa | HGIN/LP |
| B2 |  | Grossly abnormal IPCLs. Linear vessel formation with loss of the normal loop structure. There is also gross dilation and tortuosity of vessels. Formation of avascular areas between vessels observed | MM/SM1 |
| B3 |  | Dilated non-looping vessels typically 3x calibre of B2 vessels seen. IPCLs are tortuous and highly irregular in appearance. Evidence of extensive neovascularisation and avascular areas seen. Mucosal contour may be distorted. Other adjacent IPCLs are typically abnormal | SM2 or deeper |

United European Gastroenterology Journal

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

| IPCL classification | Neoplastic | Typical histology | Amenable to EET? |
| --- | --- | --- | --- |
| A | No | Normal/LGD | Yes |
| B1 | Yes | HGD/LP | Yes |
| B2 | Yes | MM/SM1 | Possibly |
| B3 | Yes | SM2 or deeper | No |

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

**Image acquisition**

ME-NBI images

Up to 80x magnification

Quality controlled

**Input image**

**Convolutional neural network (hidden layers)**

Filters applied to input images.

Each layer looks at sequentially higher resolution images

Images from patients used for training stages of network were not used for the testing phase

**eCAMs**

eCAMs generated for each of the layers within the CNN. Visual display of discriminative features the CNN used to make classification

**Output layer**

CNN makes a prediction between 0-1 for whether image contains normal or abnormal IPCL patterns

**Testing**

Computer predictions of IPCL patterns compared with gold standard (expert consenus and histology)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
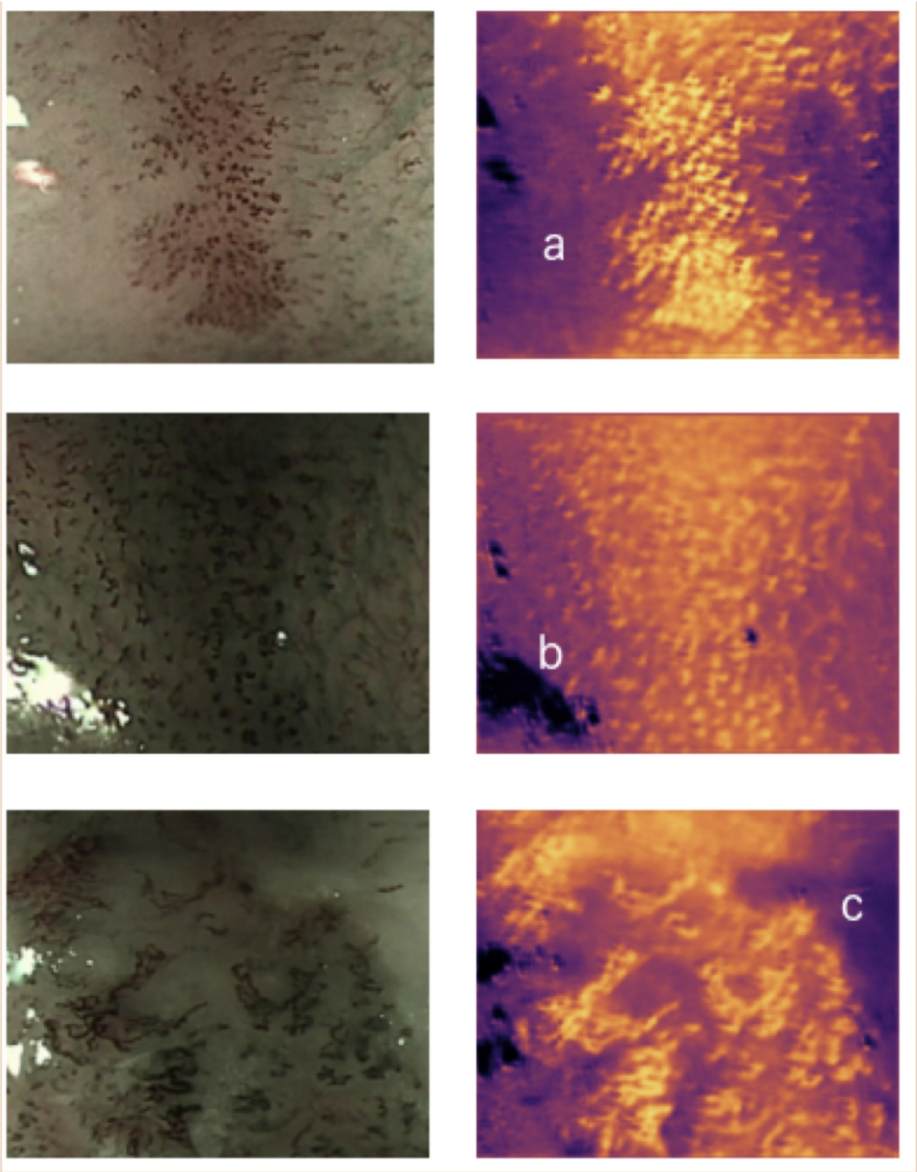48
49
50
51
52
53
54
55
56
57
58
59
60

**Figure 1a:** Representative examples of lesions containing each of the JES IPCL subtypes, with a description of key features and the associated histological invasion depth which correlates with each type. LGIN – low grade intraepithelial neoplasia, HGIN – high grade intraepithelial neoplasia, LP – lamina propria, MM – muscularis mucosa, SM – submucosa

*Figure 1b: Summary of the Japanese Endoscopic Society magnification endoscopy IPCL classification system; typical invasion depths of ESCN compared with the observed IPCL patterns. LGIN – low grade intraepithelial neoplasia. HGIN – high grade intraepithelial neoplasia. LP – lamina propria. MM – muscularis mucosa. SM - submucosa*

*Figure 2a: Representative example of how 5 fold cross validation was applied to the datasets used in of our study*

*Figure 2b: Schematic representation of our study workflow and CNN design*

**Figure 3:** Input images (left column) with corresponding eCAMs (right column), illustrating visual features recognised by the CNN when classifying images. a) recognition of abnormal IPCLs patterns. b) specular reflections are ignored by the CNN c) high selectivity between normal mucosa and abnormal IPCLs.

| Fold | Training | Validation | Testing |
|---|---|---|---|
| 1 | 2620 | 201 | 577 |
| 2 | 1792 | 891 | 715 |
| 3 | 1822 | 685 | 891 |
| 4 | 1792 | 715 | 891 |
| 5 | 1559 | 685 | 1154 |
| **Average** | 1917 | 635 | 646 |

**Table 1:** Number of frames in each fold used for cross validation containing normal IPCL patterns

| Fold | Training | Validation | Testing |
|---|---|---|---|
| 1 | 2803 | 258 | 587 |
| 2 | 2205 | 739 | 704 |
| 3 | 1549 | 1360 | 739 |
| 4 | 1912 | 961 | 775 |
| 5 | 1754 | 743 | 1151 |
| **Average** | 2045 | 812 | 791 |

**Table 2:** Number of frames in each fold used for cross validation containing abnormal IPCL patterns

| Patient demographic | | |
|---|---|---|
| ESCN characteristics | | |
| IPCL patterns | Type A | 7 |
| | Type B1 | 5 |
| | Type B2 | 5 |
| Histology | Normal | 7 |
| | HGIN (M1) | 1 |
| | Lamina propria (M2) | 4 |
| | Muscularis mucosa (M3) | 4 |
| | Submucosa (SM1) | 1 |

**Table 3:** Summary of demographics and lesion information for patients recruited

| Fold | Accuracy (%) | Sensitivity (%) | Specificity (%) | F1 score (%) |
|---|---|---|---|---|
| 1 | 86.2 | 80.4 | 92 | 85.4 |
| 2 | 89.0 | 78.1 | 99.7 | 87.6 |
| 3 | 97.7 | 100 | 95.9 | 97.6 |
| 4 | 98.3 | 99.4 | 97.3 | 98.2 |
| 5 | 95.1 | 90.6 | 99.6 | 94.9 |
| **Average** | 93.3 | 89.7 | 96.9 | 92.7 |

**Table 4:** Summary of CNN performance statistics for detection of abnormal IPCL patterns