Automated computer aided endoscopic diagnosis of early oesophageal squamous cell carcinoma using intrapapillary capillary loop patterns

M Everson[1,2] LCGP Herrera[3], W Li, I Muntion Luengo[3], O Ahmad[1,3], M Banks [1,2], C Magee[1,2], D Alzoubaidi[1,2], D Graham[1,2], T Vercauteren[3], L Lovat [1,2,3], S Ourselin[3], Hsiu-Po Wang[4], *Wen-Lun Wang[5], *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]Department of Internal Medicine, National Taiwan University, Taipei, Taiwan; [5] Department of Internal Medicine, E-Da Hospital/I-Shou University, Kaohsiung, Taiwan

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

Abstract

Background

Intrapapillary capillary loops (IPCLs) represent an endoscopically visible feature of early squamous cell neoplasia (ESCN), which correlate with lesion invasion depth. Accurate real time endoscopic assessment of depth of invasion of ESCN is vital to support decision making for local curative endoscopic therapy. Several classifications exist for use with magnification endoscopy (ME) based on IPCL morphology to determine invasion depth. IPCL patterns visualised on ME with narrow band imaging (ME-NBI) can be used to train convolutional neural networks (CNNs) to detect the presence of 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, using 5-fold cross validation. 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. Images used in training datasets were not used in validation or testing datasets.

Results

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

Conclusion

Our novel application of computer aided endoscopic diagnosis (CAED) and artificial intelligence shows that such a system can accurately classify IPCL patterns as normal or abnormal using the JES classification. Such a system could be used as an in-vivo, real-time clinical decision support tool for endoscopists undertaking endoscopic assessment and directing local therapy of ESCN.

Summary

What is already known about this subject?

Intrapapillary capillary loop (IPCL) patterns are an endoscopically visible microvascular feature of early squamous cell neoplasias (ESCN) of the oesophagus, and have been well validated as corresponding to the grade and invasion depth of neoplasia. Non-expert clinicians or those with less experience managing ESCN may struggle to accurately classify Convolutional neural networks (CNN) can be trained to recognise and classify images obtained during endoscopy. No such system currently exists for the accurate classification of IPCL patterns

What are the new findings?

We present the first CNN capable of accurately classifying abnormal IPCL patterns as normal or abnormal to assist in the in-vivo detection of early squamous cell neoplasia. To train our CNN and achieve high accuracies, we have formed the largest dataset of videos and still images of ESCNs reported in the literature used to train a CNN. We have developed a novel neural network architecture which utilises embedded class activation maps (eCAMs), to produce a clinically interpretable output of the CNN. Our system is the first reported for this application to operate in real-time, with endoscopic videos as input data.

How might it impact on clinical practice in the foreseeable future?

This platform provides proof of concept that a real-time artificial intelligence system can be developed for the detection, characterisation and prediction of ESCN invasion depth. A validated CNN, trained with a large dataset could be used clinically as a decision support tool for clinicians undertaking the assessment of ESCN.
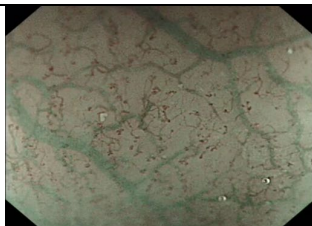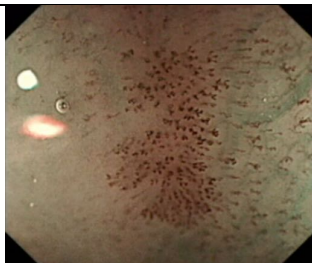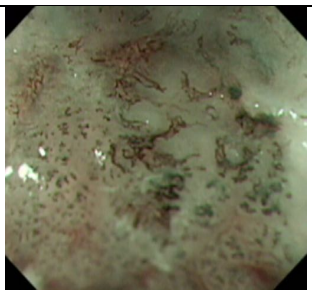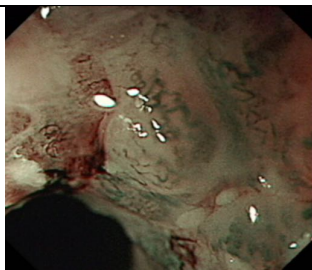
Introduction

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, through the Middle East and onwards to China and Japan[1–3]. Gastroscopy remains the investigation of choice for the diagnosis of early squamous cell neoplasia (ESCN) of the oesophagus. The endoscopic features of these early lesions are focal, subtle and so are easily missed during endoscopy - several studies have shown a significant miss rate for UGI cancers on endoscopies undertaken with three years of diagnosis[4]. Early detection of ESCN and prompt access to endoscopic eradication therapy (EET) is vital to ensure a favourable prognosis and to spare patients the attendant morbidity and mortality of oesophagectomy after a late diagnosis[5,6].

The 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 minimal invasive endoscopic therapy[7,8,9,10]. Endoscopic resection, through endoscopic mucosal resection (EMR) or submucosal dissection (ESD) is associated with excellent rates of 5 year survival[11].

Intrapapillary capillary loops (IPCLs) are microvessels that were first characterised with the advent of magnification endoscopy[12,13]. Subsequent work has established them as a marker of ESCN lesions and changes in their morphology correlates with the invasion depth of ESCN[14,15]. Normal IPCLs arise obliquely 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 a stepwise destruction of the oesophageal wall architecture, which is visibly manifested in morphologic changes in the ICPL pattern. Initially IPCLs become more tortuous and dilated as they become engorged. As the ESCN progresses and begins to interrupt the mucosal layer, the IPCLs lose their looped structure and appear as 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].

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

The recently published Japanese Endoscopic Society (JES) magnified endoscopy classification for IPCLs is a simplified system compared to early classifications, which allows the easy recognition of ESCN by endoscopists[21,15]. The JES classification has become widely accepted and used by most practitioners 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). Type A vessels are small, narrow calibre looped vessels found within normal tissue. Type B1 vessels retain their looped structure but are more tortuous and dilated and may be associated with a brown discoloration of the mucosa; indicating either high grade dysplasia (HGD) or lamina propria (LP) invasion of ESCN. Type B2 vessels are severely irregular, dilated and are associated with loss of their normal looped arrangement and correspond with muscularis mucosa (MM) or early submucosal invasion (SM1). Type B3 vessels are associated with advanced submucosal invasion or deeper and are a more aberrant formation of severely dilated B2 vessels. 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 accuracy 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].

JES type	V1	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/LGD
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	HGD/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

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. LGD – low grade dysplasia, HGD – high grade dysplasia, LP – lamina propria, MM – muscularis mucosa, SM - submucosa

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

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*

Computer aided endoscopic diagnosis with artificial intelligence 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, which are well validated as corresponding to histologic findings and invasion depth, could provide the input data to train a CNN to classify such microvasculature patterns as normal or abnormal. A CNN capable of accurately classifying the oesophageal mucosa as normal or abnormal based on IPCL patterns could improve and automate the detection and recognition of invasion depth of ESCN. This would be particularly useful in settings where endoscopists are unfamiliar with the endoscopic appearance of early squamous lesions. Furthermore, such a system would enable the in-vivo assessment of ESCNs, allowing endoscopists to direct appropriate endoscopic therapy where required and to potentially facilitate in the endoscopic delineation of resection margins.

In addition, a downstream application of such a validated system, trained to differentiate individual 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 ascertained from the IPCL patterns. Such a system could reduce procedure time and prevent endoscopic resection being used in patients for whom it would be either inappropriate or futile.

We aimed to develop a CNN capable of the automated classification of oesophageal mucosal tissue visualised using ME-NBI as non-neoplastic (type A) or neoplastic (type B1/2) based on the JES IPCL system. We chose only to include lesions containing type B1 or B2 IPCL patterns, since only these lesions would typically be deemed endoscopically resectable according to the JES classification. Our CNN was designed to work in real-time to facilitate its downstream application for the in-vivo detection and classification of ESCNs. Importantly we aimed to use images that were minimally pre-processed, in order to train a network with images typically seen at endoscopy.

Methods

Patient recruitment, inclusion and exclusion criteria

Patients attending two specialist high volume referral centre for ESCN in Taiwan |(National Taiwan University Hospital and -Da Hospital/I-Shou University) were recruited to this study. Patients were required to give valid consent. Patients with ESCN were only included if pathological samples (EMR or ESD or biopsy for invasive cancer or oesophagectomy specimens) were acquired at the time of their endoscopy or subsequent surgery. Patients with active oesophageal ulceration or anatomical abnormalities of the oesophagus were excluded. Our study complied with the Declaration of Helsinki. The Institutional Review Board of the E-Da Hospital approved this study to collect the videos (IRB number: EMRP-097-022).

Endoscopic procedures and image acquisition

Gastroscopies were performed by two expert endoscopists (WLW, HPW), defined as those performing >50 ESCN assessments per year). Endoscopies were performed using a high definition ME-NBI GIF-H260Z endoscope. An Olympus Lucera CV-290 processor (Olympus, Tokyo, Japan) was used. The oesophageal mucosa was cleaned with a solution of Simethicone prior to the interrogation of the IPCL pattern using 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 (WLW, RJH and HPW) with extensive experience in the endoscopic assessment and treatment of ESCN lesions. For each video, the visualised IPCL patterns were classified by consensus based on the JES classification system. Classifications were then correlated with histology taken from the imaged area before a final classification was assigned to each group of images. Type A IPCLs were considered normal, type B1, 2 and 3 IPCLs were considered abnormal and indicative of the previously described grade of neoplasia. Videos were sampled at 30fps to generate sequential still images which were stored as the lossless .png format.

Image quality control

Images were manually quality controlled by a senior clinician study member, with those that were blurred, contained no visible IPCLs or where the mucosa was obscured by blood or mucus deemed uninformative and removed from both the training and testing datasets. Images were cropped to remove all black borders and any identifiable or discriminative patient or clinician demography labels. This prevented the CNN identifying discriminative features on the images to distinguish between normal and abnormal cases that were not related to the IPCL patterns.

Formation of image datasets

The full dataset consisted of 7046 images, with resolutions ranging from 458x308 to 696x308 pixels. We employed 5-fold cross validation to generate five distinct datasets with different combinations of images. For any given fold, images of patients used within the training dataset were not used in the validation or testing dataset. On average, each fold used 3962 images for training, and 1637 unseen images (846 normal and 791 abnormal) for testing. The composition of the datasets is summarised in table 1 and 2.

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

Convolutional neural network
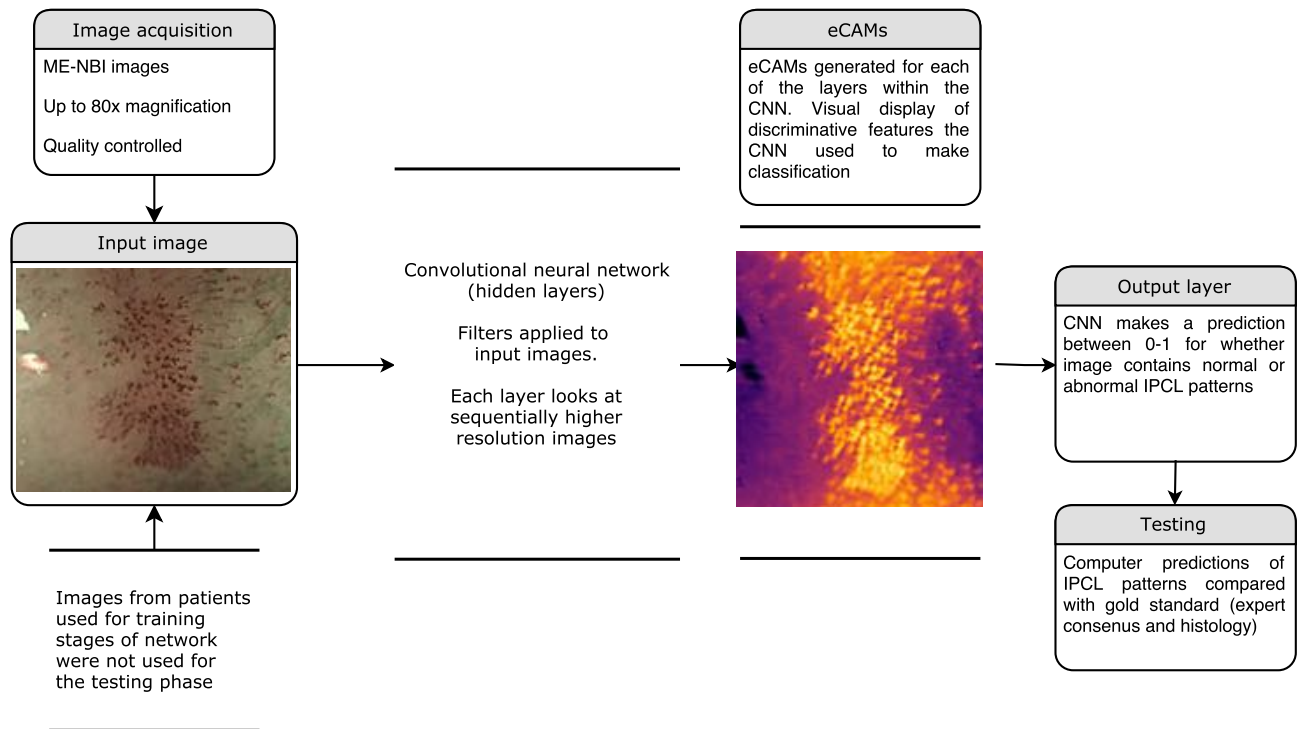
Figure 2: *Schematic representation of our study workflow and CNN design*

A full description of our technical methodology is described by our group in Herrera et al (cite Luis), a summary is provided in figure 2. 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 are passed through layers of the CNN, which develops filters for features on the image, at various magnifications. As images pass through the layers of the CNN, the network develops filters to detect various visual features such as edges and borders, colours and the shape of IPCL patterns. 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. In a final output layer the CNN will make a prediction whether the IPCL pattern is normal or abnormal. If the prediction is correct compared to the clinician labelled input image, those filters will be 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 detection were calculated.

Results

Patient characteristics

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

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

Analysis of explicit class activation maps (eCAMs)

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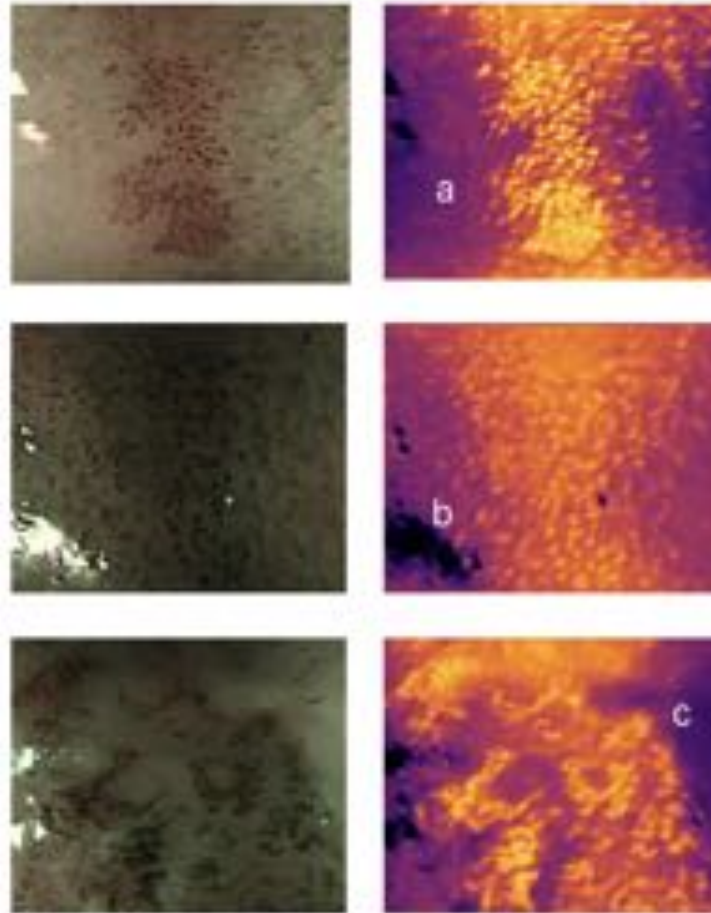
Figure 3: Input images (left column) with corresponding eCAMs (right column), illustrating visual features recognised by the CNN when classifying images. a) recognition of abnormal IPCLs patterns. b) specular reflections are ignored by the CNN c) high selectivity between normal mucosa and abnormal IPCLs.

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 detection 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 fully train the network to recognise the full spectrum of variability in IPCL patterns. Our CNN performance statistics are summarised in table 4.

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

Discussion

We have introduced the first application of CAED and artificial intelligence for the accurate, real-time characterisation of ESCN through the classification of intrapapillary capillary loops seen during magnification endoscopy based on the widely used clinical endoscopic 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%) compared to lesions invading the submucosa (8-45.9%) and so are amenable to minimal invasive endoscopic therapy[7,8,9,10]. Prompt access to endoscopic eradication therapies and the low incidence of metastasis, mean that superficial and mucosal ESCN can be resected with 5 year survivals of 75-100%[11].

Several classification systems exist to assist clinicians with the accurate characterisation of IPCL patterns. Inoue et al (2001) were the very first to propose a five-part classification of IPCL patterns[25]. Type I-III IPCL patterns were associated with normal mucosa, inflammation or LGIN. Type IV IPCLs were suggestive of HGD. Type V IPCL patterns 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[25]. 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]. The Arima classification also introduced the concept of avascular areas as a marker of invasive cancer, with AVAs <5mm suggestive of SM1 invasion and >5mm of >SM2 invasion[20].

We elected to use the JES IPCL classification of ESCN as the input and output for our neural network as it is the most contemporary and, as described above, accurate classification system that has been proposed and is widely used in areas of high ESCN prevalence. We believe that a CNN that functions using this classification will remain clinically relevant as it becomes more widely used. We suggest that the classifications proposed by Inoue[25] and Arima[20] are complex and may not be as intuitive to use in clinical practice.

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 also compares favourably to other published work at 89.7% and 96.9% respectively. Furthermore, our CNN can operate at video rate, with rapid prediction times meaning that it could be used for real-time in vivo classification of imaged mucosal tissue.

We envisage that our CNN could provide a 'red-flag' type system to assist endoscopists in the recognition of abnormal IPCL patterns in patients undergoing assessment for suspected ESCN and consideration of endoscopic treatment. Such a validated system could potentially 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[26]. 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 yield most benefit for non-expert endoscopists or in 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 which we have tried to mitigate; further work currently underway seeks to address them too. Firstly, our patient sample size is small, using 7046 images generated from 17 patients. Although we note that 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 generalise – defined as its ability 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 that were used for training were used for testing. Our novel use of eCAMs also acts as an additional form of validation; despite the low total patient numbers, we have already demonstrated that the ICPL patterns are the visual features the network finds discriminative when classifying images. We therefore expect that with greater numbers of patients we should continue to see such high accuracies for IPCL characterisation.

Secondly, our system is able to differentiate normal (type A) compared to abnormal (type B1-3) IPCL patterns, but is not yet able to differentiate between 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 can highlight abnormal areas of potentially resectable lesions to clinicians in the first instance. A system capable of identifying type B3 lesions, which have typically invaded beyond SM2 and so are currently

not typically endoscopically resectable requires further work and robust validation before it could be used routinely.

Lastly, our gold standard used only three expert clinicians. We feel that for such a pilot study this is sufficient to provide accurate ground truth to develop a CNN, particularly since all of the consensus classifications were correlated with histological results. We recognise however that future work will be needed to externally validate this CNNs performance against a larger panel of clinicians of varying levels of 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.

Based on further work currently being undertaken we aim to produce a CNN capable of more precise classification based on the individual JES subgroups – type A, B1, B2 and B3. Such a validated, accurate system would also improve the characterisation of ESCN. Furthermore, this system could allow clinicians to triage, in real—time, ESCN lesions that would be amenable to EET, compared to those lesions which would be too advanced to be cured endoscopically. This would both prevent patients with advanced disease undergoing unnecessary endoscopic therapy, as well as ensuring that they were referred promptly for surgical management. Such a system would need to be validated in a prospective clinical trial before it could be adopted into routing clinical practice. Work already underway by our group aims to develop a segmentation tool which could be used in real time to delineate borders of neoplastic lesions to ensure adequate resection margins.

In conclusion, our pilot study introduces a novel CNN, 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. We believe that our system provides a performance benchmark for the assessment of other CNNs designed for this purpose in the future.

Word count 3538 – excluding abstracts, figures, references

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. 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).

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

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

8. 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).

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

10. 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).

11. 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).

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

26.　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

27.　Wen-Lun, W. *et al.* A training program of a new simplified classification of magnified

narrow band imaging for superficial esophageal squamous cell carcinoma. *J. Gastroenterol. Hepatol.* **0,** (2017).