

# The Application of Deep Learning in Prognosis Prediction for Solid Tumor Imaging

Lead Guest Editor: Hao Wang

Guest Editors: Hong-Ning Dai and Liwei Hu





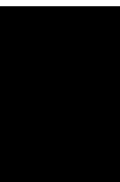
---

# **The Application of Deep Learning in Prognosis Prediction for Solid Tumor Imaging**

# **The Application of Deep Learning in Prognosis Prediction for Solid Tumor Imaging**

Lead Guest Editor: Hao Wang

Guest Editors: Hong-Ning Dai and Liwei Hu



---

Copyright © 2023 Hindawi Limited. All rights reserved.

This is a special issue published in "Journal of Healthcare Engineering." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Associate Editors

Xiao-Jun Chen , China  
Feng-Huei Lin , Taiwan  
Maria Lindén, Sweden

## Academic Editors

Cherif Adnen, Tunisia  
Saverio Affatato , Italy  
Óscar Belmonte Fernández, Spain  
Sweta Bhattacharya , India  
Prabadevi Boopathy , India  
Weiwei Cai, USA  
Gin-Shin Chen , Taiwan  
Hongwei Chen, USA  
Daniel H.K. Chow, Hong Kong  
Gianluca Ciardelli , Italy  
Olawande Daramola, South Africa  
Elena De Momi, Italy  
Costantino Del Gaudio , Italy  
Ayush Dogra , India  
Luobing Dong, China  
Daniel Espino , United Kingdom  
Sadiq Fareed , China  
Mostafa Fatemi, USA  
Jesus Favela , Mexico  
Jesus Fontecha , Spain  
Agostino Forestiero , Italy  
Jean-Luc Gennisson, France  
Badicu Georgian , Romania  
Mehdi Gheisari , China  
Luca Giancardo , USA  
Antonio Gloria , Italy  
Kheng Lim Goh , Singapore  
Carlos Gómez , Spain  
Philippe Gorce, France  
Vincenzo Guarino , Italy  
Muhammet Gul, Turkey  
Valentina Hartwig , Italy  
David Hewson , United Kingdom  
Yan Chai Hum, Malaysia  
Ernesto Iadanza , Italy  
Cosimo Ieracitano, Italy

Giovanni Improta , Italy  
Norio Iriguchi , Japan  
Mihajlo Jakovljevic , Japan  
Rutvij Jhaveri, India  
Yizhang Jiang , China  
Zhongwei Jiang , Japan  
Rajesh Kaluri , India  
Venkatachalam Kandasamy , Czech Republic  
Pushpendu Kar , India  
Rashed Karim , United Kingdom  
Pasi A. Karjalainen , Finland  
John S. Katsanis, Greece  
Smith Khare , United Kingdom  
Terry K.K. Koo , USA  
Srinivas Koppu, India  
Jui-Yang Lai , Taiwan  
Kuruva Lakshmanna , India  
Xiang Li, USA  
Lun-De Liao, Singapore  
Qiu-Hua Lin , China  
Aiping Liu , China  
Zufu Lu , Australia  
Basem M. ElHalawany , Egypt  
Praveen Kumar Reddy Maddikunta , India  
Ilias Maglogiannis, Greece  
Saverio Maietta , Italy  
M.Sabarimalai Manikandan, India  
Mehran Moazen , United Kingdom  
Senthilkumar Mohan, India  
Sanjay Mohapatra, India  
Rafael Morales , Spain  
Mehrbakhsh Nilashi , Malaysia  
Sharnil Pandya, India  
Jialin Peng , China  
Vincenzo Positano , Italy  
Saeed Mian Qaisar , Saudi Arabia  
Alessandro Ramalli , Italy  
Alessandro Reali , Italy  
Vito Ricotta, Italy  
Jose Joaquin Rieta , Spain  
Emanuele Rizzuto , Italy

Dinesh Rokaya, Thailand  
Sébastien Roth, France  
Simo Saarakkala , Finland  
Mangal Sain , Republic of Korea  
Nadeem Sarwar, Pakistan  
Emiliano Schena , Italy  
Prof. Asadullah Shaikh, Saudi Arabia  
Jiann-Shing Shieh , Taiwan  
Tiago H. Silva , Portugal  
Sharan Srinivas , USA  
Kathiravan Srinivasan , India  
Neelakandan Subramani, India  
Le Sun, China  
Fabrizio Taffoni , Italy  
Jinshan Tang, USA  
Ioannis G. Tollis, Greece  
Ikram Ud Din, Pakistan  
Sathishkumar V E , Republic of Korea  
Cesare F. Valenti , Italy  
Qiang Wang, China  
Uche Wejinya, USA  
Yuxiang Wu , China  
Ying Yang , United Kingdom  
Elisabetta Zanetti , Italy  
Haihong Zhang, Singapore  
Ping Zhou , USA

## Contents








---

**Differentiation of Small Gastrointestinal Stromal Tumor and Gastric Leiomyoma with Contrast-Enhanced CT**

Mingyan Yan , Yubao Liu , Honglian You , Yanbo Zhao , Jun Jin , and Jing Wang 








Research Article (6 pages), Article ID 6423617, Volume 2023 (2023)

**A Novel Framework for Melanoma Lesion Segmentation Using Multiparallel Depthwise Separable and Dilated Convolutions with Swish Activations**

Maryam Bukhari , Sadaf Yasmin , Adnan Habib , Xiaochun Cheng , Farhan Ullah , Jaeseok Yoo , and Daewon Lee 

Research Article (15 pages), Article ID 1847115, Volume 2023 (2023)

**Evaluating Histological Subtypes Classification of Primary Lung Cancers on Unenhanced Computed Tomography Based on Random Forest Model**

Jianfeng Huang , Wei He , Haijia Xu , Shan Yang , Jiajun Dai , Weifeng Guo , and Mengsu Zeng 

Research Article (8 pages), Article ID 8964676, Volume 2023 (2023)

## Research Article

# Differentiation of Small Gastrointestinal Stromal Tumor and Gastric Leiomyoma with Contrast-Enhanced CT

Mingyan Yan <sup>1</sup>, Yubao Liu <sup>2</sup>, Honglian You <sup>1</sup>, Yanbo Zhao <sup>2</sup>, Jun Jin <sup>1</sup>,  
and Jing Wang <sup>1</sup>

<sup>1</sup>Department of Radiology, Shenzhen Qianhai Shekou Free Trade Zone Hospital, Shenzhen 518000, China

<sup>2</sup>Imaging Center, Shenzhen Hospital of Southern Medical University, Shenzhen 518100, China

Correspondence should be addressed to Jing Wang; you28154314@163.com

Received 31 May 2022; Revised 7 July 2022; Accepted 24 November 2022; Published 8 February 2023

Academic Editor: Liwei Hu

Copyright © 2023 mingyan Yan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Objective.** The value of multiphase contrast-enhanced CT in differentiating gastrointestinal stromal tumors (GISTs) and gastric leiomyomas (GLMs) which were  $\leq 3$  cm was evaluated using machine learning. **Methods.** A retrospective analysis was conducted on 45 cases of small gastric wall submucosal tumors (including 22 GISTs and 23 GLMs) with pathologically confirmed diameter  $\leq 3$  cm and completed multiphase CT-enhanced scan images. The CT features including tumor location, maximum diameter, shape, margins, growth pattern, plain/enhanced CT value, cystic degeneration, calcification, ulcer, progressive reinforcement, perilesional lymph nodes, and the CT value ratio of the tumor to the aorta at the same level in the enhanced phase III scan of the two groups were evaluated. Tumor location and maximum diameter were automatically evaluated by machine learning. **Results.** The GISTs and GLMs with a diameter  $\leq 3$  cm showed clear margins, uniform density on plain scan CT, and progressive homogeneous enhancement. The age of the GISTs is greater than that of the GLMs group. The plain scan CT value of the GISTs group was lower than that in the GLMs group. In the GISTs group, the lesions were mostly located in the fundus (68.18%), showing a mixed growth pattern (54.55%), and in the GLMs group, most lesions were located in the cardia (47.82%), showing an intraluminal growth pattern (95.65%). The abovementioned differences were statistically significant. **Conclusions.** Contrast-enhanced CT has limited value in differentiating small GISTs from GLMs, which are  $\leq 3$  cm. Older age ( $>49.0$  years), a low plain CT value ( $<42.5$  Hu), mixed growth inside and outside the cavity, and noncardiac location tended to be the criteria for the diagnosis of small GISTs of the gastric wall.

## 1. Introduction

Gastrointestinal stromal tumors (GISTs) and gastric leiomyomas (GLMs) are two common tumors derived from gastrointestinal mesenchymal tissue. GISTs are Cajal cells originating from the muscularis propria of the gastrointestinal wall and have malignant potential [1], while GLMs are benign tumors originating from the smooth muscle tissue of the gastrointestinal wall [2]. GIST lesions vary in size, and their imaging manifestations are diverse, with round, quasi-round, or irregular shapes. Necrotic cystic degeneration, hemorrhage, and calcification may occur in the lesions. Enhanced scanning shows uniform or uneven enhancement, most of which are

obvious, and vascular-like enhancement can be seen in some arterial phases, while necrosis and cystic degeneration are not significantly enhanced. GLMs mostly occur in the stomach. In addition to direct invasion and distant metastasis suggesting malignancy, irregular or lobulated tumors, uneven enhancement, central necrosis, ulceration, and uneven thickening of the adjacent intestinal wall suggest a high possibility of LMs. Submucosal tumors of the stomach wall with diameter  $\leq 3$  cm in the GISTs are GISTs with a decreasing diameter, and the value of multiphase CT enhancement in the identification of gastric wall GISTs and GLMs is limited. Among them, only older age, a lower plain CT value, mixed growth inside and outside the cavity, and occurrence in the gastric



fundus are more inclined to the diagnosis of small GISTs in the gastric wall.

Previous studies on the differential diagnosis of the two tumors in imaging are not lacking. With the increasing widespread clinical application of digestive endoscopy, more and more small gastric wall submucosal tumors have been discovered and attracted attention [3]. As the tumor diameter decreases, its imaging features are incharacteristic, which makes differential diagnosis more difficult, and preoperative differentiation between small GISTs and GLMs is important for treatment selection. In this study, multiphase enhanced CT images of gastric wall submucosal tumors with a diameter of  $\leq 3$  cm were retrospectively analyzed to explore their imaging manifestations and the basis for differential diagnosis.

## 2. Materials and Methods

**2.1. Research Object.** A total of 78 cases of pathologically diagnosed gastric wall submucosal tumors were treated by endoscopic ultrasonic-guided fine-needle puncture or endoscopic/surgical resection from June 2018 to June 2021. Among them, there were 35 cases of GISTs, 31 cases of GLMs, 5 cases of gastric schwannoma, 4 cases of ectopic pancreas, 1 case of gastric polyp, 1 case of gastric cancer, and 1 case of glomus tumor. GISTs and GLMs with a diameter of  $\leq 3$  cm were included in this study, and cases without preoperative CT enhancement or poor image quality were excluded. Finally, 45 cases were included, including 22 cases of GISTs and 23 cases of GLMs.

**2.2. Inspection Method.** All examinations were performed with a 256-slice multislice CT scanner (Revolution CT, GE, USA). Before the CT examination, the patient fasted for 4–6 hours, and 500–1000 mL of water was taken orally 10 minutes before the scan (to fully dilate the stomach and duodenum). Scanning parameters were as follows: tube voltage 120 kV, Smart mA technology, speed 0.5 s, pitch 1.375, layer thickness 5 mm. The reconstruction layer thickness is 1.25 mm. The patient underwent breathing training. After the plain scan, a high-pressure syringe was used to inject a nonionic iodinated contrast agent (iohexol, 300 mgI/mL, 1.2 ml/kg) through the antecubital vein at a flow rate of 2.5–3.5 ml/s, and the scan was triggered by the smart tracking technology. The abdominal aorta layer was monitored; the trigger threshold was 120 Hu. The breathing command was deep inhalation and then breath-holding, and multiphase CT-enhanced scanning was performed in the arterial phase (5.9 s), the portal venous phase (20 s), and the delayed phase (120 s) by automatically triggering the enhanced scanning. The tube was flushed with 30 ml of normal saline after the injection.

**2.3. Research Method.** The blind method is used. The CT images were retrospectively analyzed by two imaging diagnosticians (with working years of 15 and 7 years, respectively), and the differences were resolved through consultation. The contents of analysis include tumor

location, maximum diameter, shape, margins, growth pattern (intraluminal or mixed type, this group of cases is small, and there is no single extraluminal growth case), plain scan and enhanced CT values at each stage (select the largest level of the lesion, the area of uniform density, the area of interest  $2\sim 4$  mm<sup>2</sup>, the average CT value), the presence or absence of cystic calcification, enhancement mode (the difference between the CT value of the most obvious part of the lesion and the weakest part of the enhancement is less than 10 Hu defined as uniform enhancement, greater than or equal to 10 Hu for heterogeneous enhancement), lymph nodes around the lesion (shorter diameter greater than 5 mm), and the ratio of the three-phase enhancement of the lesion to the CT value of the aorta at the same level (R-A, R-V, R-D). When measuring aortic CT values, the ROI was enlarged as much as possible, and the average value was taken. Tumor location and maximum diameters were automatically evaluated using machine learning by GE software (GE Milwaukee, USA).

**2.4. Statistical Analysis.** SPSS 26.0 statistical software was used. Measurement data (age, maximum diameter of tumor, plain/enhanced CT value, and R-A/V/D) between two groups were compared by a *t*-test or Mann–Whitney *U* test. Data were expressed as mean  $\pm$  standard deviation ( $\bar{x} \pm s$ ) or median (upper and lower interquartile range: M (P25, P75)). A chi-square test or adjusted chi-square test was used to compare count data (sex, tumor location, morphological margin, growth pattern, presence or absence of cystic calcification, and enhancement pattern) between the two groups. The receiver operating characteristic (ROC) curve was used to determine the best cutoff value (the sum of specificity and sensitivity was the highest) to distinguish GISTs from GLMs, Youden value = sensitivity – (1 – specificity).  $P < 0.05$  was statistically significant.

## 3. Results

**3.1. Comparison of General Clinical Data and CT Features.** A total of 45 patients were included in this study, including 20 males and 25 females, aged from 27 to 68 years old, with an average age of  $(51.00 \pm 11.10)$  years, and lesion sizes ranging from 0.8 to 3.0 cm. Among them, there were 22 cases of the small GISTs in the gastric wall (pathology showed extremely low risk in 15 cases, low risk in 6 cases, and medium risk in 1 case) and 23 cases of the small GLMs in the gastric wall. The average age, maximum size of the tumor, plain CT value, multistage enhanced CT value, the ratio of the three-phase enhancement of the lesion to the CT value of the aorta at the same level, tumor location, morphology, tumor margin, growth mode, cystic change, calcification, enhancement mode, and adjacent enlarged lymph nodes of the two groups are shown in Table 1. The median age of the GISTs is greater than the median age of the GLMs group ( $P < 0.05$ ). The plain CT value of GISTs was lower than that of the GLMs group ( $P < 0.05$ ). 15/22 (68.18%) of the small GISTs in the gastric wall are located in the fundus of the stomach (Figure 1), and the 11/23 (47.83%) of small GLMs

TABLE 1: Comparison of general clinical data and CT features between small gastric GISTs and GLMs ( $n$  (%)).

	GISTs group (22 cases)	GLMs group (23 cases)	Statistical	$P$ value
Gender			1.779 <sup>①</sup>	0.182
Male	12 (54.54)	8 (34.78)		
Female	10 (45.46)	15 (65.22)		
Average age (years)	60.50 (51.00, 64.25)	47.00 (37.00, 54.00)	-3.863 <sup>④</sup>	<0.001*
Tumor location			13.276 <sup>①</sup>	0.001*
Cardia	2 (9.09)	11 (47.83)		
Fundus of stomach	15 (68.18)	4 (17.39)		
Gastric body	5 (22.73)	8 (34.78)		
Tumor shape			0.987 <sup>②</sup>	1.000
Smooth	22 (100)	22 (95.65)		
Lobular	0 (0)	1 (4.35)		
Tumor margin			2.222 <sup>②</sup>	0.243
Clear	20 (90.91)	17 (73.91)		
Verge	2 (9.09)	6 (26.09)		
Growth mode			13.792 <sup>①</sup>	<0.001*
Intraluminal	10 (45.45)	22 (95.65)		
Mixed	12 (54.55)	1 (4.35)		
Cystic change			0.301 <sup>②</sup>	0.489
Yes	1 (4.76)	0 (0)		
No	21 (95.24)	23 (100)		
Calcification			0.001 <sup>②</sup>	1.000
Yes	1 (4.76)	1 (4.35)		
No	21 (95.24)	22 (95.65)		
Enhancement mode			0.178 <sup>②</sup>	1.000
Homogeneous	20 (90.91)	20 (86.96)		
Inhomogeneous	2 (9.09)	3 (13.04)		
Progressive enhancement			0.650 <sup>①</sup>	0.420
Yes	15 (68.18)	13 (56.52)		
No	7 (31.82)	10 (43.48)		
Adjacent enlarged lymph nodes			2.001 <sup>②</sup>	0.489
Yes	0 (0)	2 (8.70)		
No	22 (100)	21 (91.30)		
Tumor maximum diameter (cm)	1.80 (1.40, 2.33)	1.30 (1.10, 1.90)	-1.707 <sup>④</sup>	0.088
Plain CT value (Hu)	37.32 ± 7.82	46.48 ± 8.63	-3.727 <sup>③</sup>	0.001*
CT value of phase A (Hu)	57.50 ± 12.31	59.65 ± 8.93	-0.674 <sup>③</sup>	0.504
R-A	0.19 (0.15, 0.21)	0.18 (0.17, 0.21)	-0.352 <sup>③</sup>	0.725
CT value of phase V (Hu)	69.86 ± 11.74	67.43 ± 9.63	0.760 <sup>③</sup>	0.451
R-V	0.49 (0.45, 0.55)	0.44 (0.42, 0.52)	-1.953 <sup>④</sup>	0.051
CT value of phase D (Hu)	73.23 ± 11.95	73.43 ± 10.05	-0.063 <sup>③</sup>	0.950
R-D	0.71 ± 0.10	0.68 ± 0.09	1.140 <sup>③</sup>	0.261

Note. ① represents the chi-square test, ② represents the correction chi-square test, ③ represents the  $T$  test, ④ represents the Mann-Whitney  $U$  test; \*indicates a statistical difference.

in the stomach wall are located in the cardia (Figure 2). The small GISTs in the stomach wall with mixed growth in the intraluminal/extraluminal are about 12/22 (54.55%) (Figure 1), whereas the GLMs are mainly with intraluminal growth (Figure 2) and are about 22/23 (95.65%) ( $P < 0.05$ ).

**3.2. Receiver Operating Characteristic (ROC) Curve.** The ROC curve showed that when patients were older than 49.50 years old, the area under the curve, sensitivity, and specificity for diagnosing GISTs were 0.836, 86.4%, and 73.9%, respectively (Figure 3). When the unenhanced CT value was greater than 42.5 Hu, the area under the curve, sensitivity, and specificity for diagnosing GLMs were 0.777, 69.6%, and 77.3%, respectively (Figure 4).

## 4. Discussion

With the wide application of digestive endoscopy, the detection of gastric subepithelial lesions has greatly increased, especially the small GISTs in the gastric wall with a diameter of  $\leq 3$  cm. Due to its inert biological characteristics, some scholars suggest that it should be classified as a special type or even a benign tumor to be distinguished. So what is the imaging performance and identification basis of such small GISTs and small GLMs? Is there any change compared with conventional lesions reported in the literature?

This study showed that the average age of the small GISTs group was greater than that of the GLMs group, which

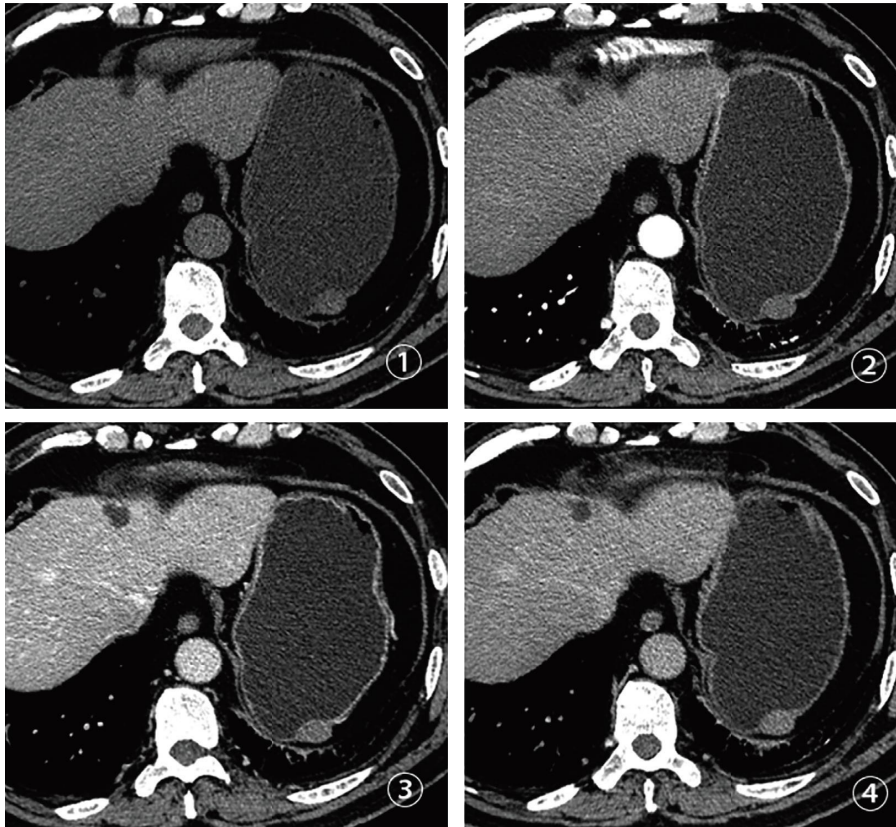


FIGURE 1: A 63-year-old man with small GISTs in the stomach wall. ① Plain CT scan revealing an oval mixed growth of soft tissue mass in the fundus of the stomach, smooth with clear margin and a maximum diameter of 2.1 cm. The CT value of plain CT scan was 32 Hu, showing mixed growth in the intraluminal/extraluminal. ②–④ Enhancement scanning of the arterial phase, venous phase, and delayed phase, resulting in uniform progressive enhancement of the mass.

was consistent with the results of previous studies [4]. The ROC curve showed that when the patient was older than 49.50 years old, the area under the curve, sensitivity, and specificity of GISTs were 0.836, 86.4%, and 73.9%, respectively, indicating that older age at onset can also be used as the basis for the identification of gastric small GISTs and small GLMs.

Choi et al. studied that the gastric corpus is the predisposing site of GISTs in the gastric wall [5], and the risk of GISTs increases as the GISTs move down in the gastric cavity compared with the gastric cardia [6]. In this study, the small GISTs were mostly located in the gastric fundus, and 95.45% were in the low-risk or very low-risk group. Cardia or involvement of the gastroesophageal junction is a significant feature of GLMs [5, 7]. In this study, about 47.82% of small GLMs occurred in the cardia. Therefore, the occurrence of lesions in the gastric fundus or the cardia becomes one of the important basis for the identification of gastric wall small GISTs and small GLMs.

The plain CT value of the GISTs group was lower than that of the GLMs group. The ROC curve showed that when the unenhanced CT value was greater than 42.5 Hu, the area under the curve, sensitivity, and specificity for diagnosing GLMs were 0.777, 69.6%, and 77.3%, respectively. In the past, the imaging characteristics of GISTs have rarely paid attention to the CT value of a plain scan, and more attention

has been paid to whether the density of the lesion is uniform and whether there is cystic necrosis and hemorrhage. These signs suggest an increased risk of malignancy and become the main basis for distinguishing GISTs from other benign tumors of the gastric wall. In our study, the small lesions had clear margins and uniform enhancement on both plain and enhanced scans, indicating the characteristics of low or very low malignancy. The lower CT value of a plain scan has become one of the bases for diagnosing small GISTs.

In terms of tumor growth patterns, the lesions in the GISTs group showed mixed growth, and the lesions in the GLMs group grew into the cavity, which may be due to the different histological sites of the two tumors. Endoscopic ultrasonography shows that gastric wall GISTs mostly occur in the fourth layer of the gastric wall, namely, the muscularis propria, so they can grow both inside and outside the cavity [6]. GLMs mostly originate from the third layer, the muscularis mucosae, so they grow toward the mucosal surface with less growth resistance and protrude into the lumen, which is consistent with the results of previous studies [7–9], indicating that the histological location and growth pattern of small GISTs are similar to conventional lesions.

In previous studies, in GISTs tumors with a diameter of 2–5 cm or larger, the cystic degeneration, lobulation, inhomogeneous enhancement, and surface ulcers suggested a higher risk of GISTs and became the main basis for

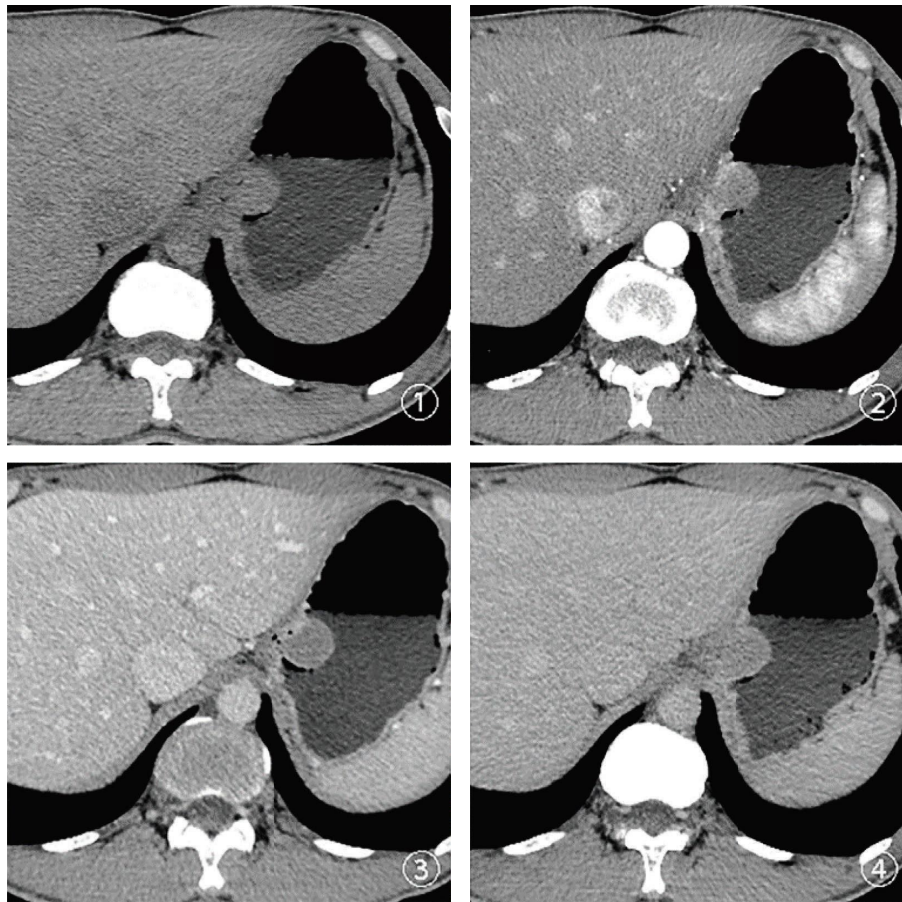


FIGURE 2: A 37-year-old man with small GLMs in the gastric wall. ① Plain CT scan showing a soft tissue mass growing in an oval cavity from the cardia, smooth with clear edge and a maximum diameter of 2.7 cm. The CT value was 38 Hu and the mass grew inside the cavity. ②–④ Enhancement scanning of the arterial phase, venous phase, and delayed phase, respectively, resulting in homogeneous enhancement of the mass.

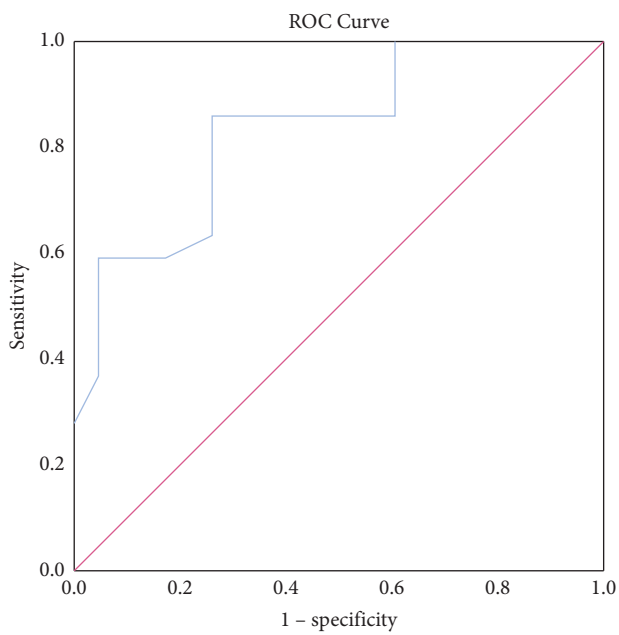


FIGURE 3: ROC curves of age-diagnosed GISTs.

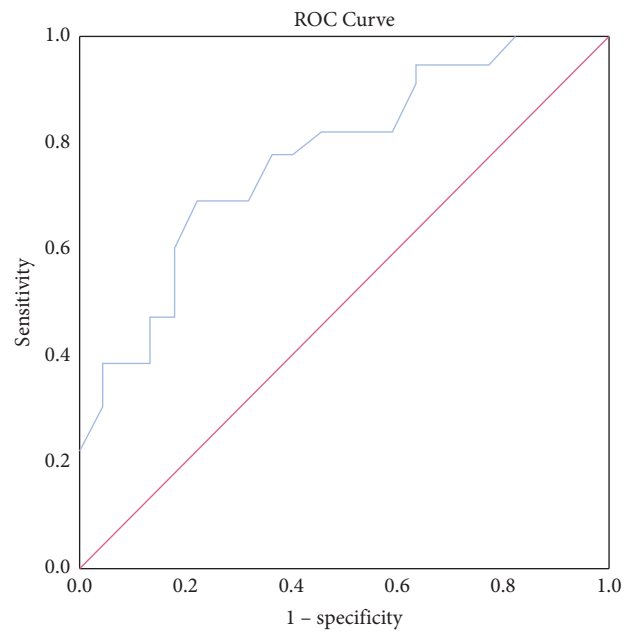


FIGURE 4: ROC curve for the diagnosis of GLMs by plain CT values.

differentiating them from GLMs [5, 10]. However, the abovementioned signs in this study of small lesions did not have any distinguishing value. Both lesions showed progressive enhancement with a clear margin and uniformity, showing the imaging manifestations of benign spindle cell-derived tumors. The disadvantage of this study is that the sample size is small, and we look forward to expanding the sample size in the future and further exploring its diagnostic and differential diagnosis characteristics through radiomics methods.

### Data Availability

The data used to support the findings of this study are available from the corresponding author upon reasonable request.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Authors' Contributions

Mingyan Yan and Yubao Liu contributed equally to this work.

### References

- [1] J. Wu, X. Huang, and Y. Zeng, "Progress in the diagnosis and treatment of gastrointestinal stromal tumors," *World Latest Medicine Information (Electronic Version)*, vol. 21, no. 15, pp. 66–69, 2021.
- [2] Y. Zheng, T. Lan, X. Huang, and W. Li, "Observation of clinicopathologic tissue and imaging features of gastric leiomyoma and gastric stromal tumors," *China Modern Medicine*, vol. 27, no. 11, pp. 7–11, 2020.
- [3] S. Liu, X. Zhou, Y. Yao, K. Shi, M. Yu, and F. Ji, "Resection of the gastric submucosal tumor (G-SMT) originating from the muscularis propria layer: comparison of efficacy, patients' tolerability, and clinical outcomes between endoscopic full-thickness resection and surgical resection," *Surgical Endoscopy*, vol. 34, no. 9, pp. 4053–4064, 2020.
- [4] X. Yin, Y. Yin, X. Liu et al., "Identification of gastrointestinal stromal tumors from leiomyomas in the esophagogastric junction: a single-center review of 136 cases," *Medicine*, vol. 99, no. 17, Article ID e19884, 2020.
- [5] Y. R. Choi, S. H. Kim, S.-A. Kim et al., "Differentiation of large ( $\geq 5$  cm) gastrointestinal stromal tumors from benign sub-epithelial tumors in the stomach: radiologists' performance using CT," *European Journal of Radiology*, vol. 83, no. 2, pp. 250–260, 2014.
- [6] R. T. Schulz, L. C. Fabio, M. C. Franco, S. A. Siqueira, P. Sakai, and F. Maluf-Filho, "Predictive features for histology of gastric subepithelial lesions," *Arquivos de Gastroenterologia*, vol. 54, no. 1, pp. 11–15, 2017.
- [7] H. K. Yang, Y. H. Kim, Y. J. Lee et al., "Leiomyomas in the gastric cardia: CT findings and differentiation from gastrointestinal stromal tumors," *European Journal of Radiology*, vol. 84, no. 9, pp. 1694–1700, 2015.
- [8] J.-X. Xu, Q.-L. Ding, Y.-F. Lu, S. F. Fan, Q. P. Rao, and R. S. Yu, "A scoring model for radiologic diagnosis of gastric leiomyomas (GLMs) with contrast-enhanced computed tomography (CE-CT): differential diagnosis from gastrointestinal stromal tumors (GISTs)," *European Journal of Radiology*, vol. 134, Article ID 109395, 2021.
- [9] S. Okanou, M. Iwamuro, T. Tanaka et al., "Scoring systems for differentiating gastrointestinal stromal tumors and schwannomas from leiomyomas in the stomach," *Medicine*, vol. 100, no. 40, Article ID e27520, 2021.
- [10] M. Liu, L. Liu, and E. Jin, "Gastric sub-epithelial tumors: identification of gastrointestinal stromal tumors using CT with a practical scoring method," *Gastric Cancer*, vol. 22, no. 4, pp. 769–777, 2019.

## Research Article

# A Novel Framework for Melanoma Lesion Segmentation Using Multiparallel Depthwise Separable and Dilated Convolutions with Swish Activations

Maryam Bukhari <sup>1</sup>, Sadaf Yasmin <sup>1</sup>, Adnan Habib <sup>2</sup>, Xiaochun Cheng <sup>3</sup>,  
Farhan Ullah <sup>4</sup>, Jaeseok Yoo <sup>5</sup> and Daewon Lee <sup>5,6</sup>

<sup>1</sup>Department of Computer Science, COMSATS University Islamabad, Attock Campus, Attock, Pakistan

<sup>2</sup>Department of Computer Engineering, UET Taxila, Taxila, Pakistan

<sup>3</sup>Department of Computer Science, Swansea University, Bay Campus, Fabian Way, Swansea SA1 8EN, Wales, UK

<sup>4</sup>School of Software, Northwestern Polytechnical University, Xian 710072, China

<sup>5</sup>Graduate School of Advanced Imaging Science, Chung-Ang University, Seoul, Republic of Korea

<sup>6</sup>School of Art and Technology, College of Art and Technology, Chung-Ang University, Seoul, Republic of Korea

Correspondence should be addressed to Daewon Lee; [dwlee@cau.ac.kr](mailto:dwlee@cau.ac.kr)

Received 30 May 2022; Revised 16 August 2022; Accepted 24 November 2022; Published 6 February 2023

Academic Editor: Liwei Hu

Copyright © 2023 Maryam Bukhari et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Skin cancer remains one of the deadliest kinds of cancer, with a survival rate of about 18–20%. Early diagnosis and segmentation of the most lethal kind of cancer, melanoma, is a challenging and critical task. To diagnose medicinal conditions of melanoma lesions, different researchers proposed automatic and traditional approaches to accurately segment the lesions. However, visual similarity among lesions and intraclass differences are very high, which leads to low-performance accuracy. Furthermore, traditional segmentation algorithms often require human inputs and cannot be utilized in automated systems. To address all of these issues, we provide an improved segmentation model based on depthwise separable convolutions that act on each spatial dimension of the image to segment the lesions. The fundamental idea behind these convolutions is to divide the feature learning steps into two simpler parts that are spatial learning of features and a step for channel combination. Besides this, we employ parallel multidilated filters to encode multiple parallel features and broaden the view of filters with dilations. Moreover, for performance evaluation, the proposed approach is evaluated on three different datasets including DermIS, DermQuest, and ISIC2016. The finding indicates that the suggested segmentation model has achieved the Dice score of 97% for DermIS and DermQuest and 94.7% for the ISBI2016 dataset, respectively.

## 1. Introduction

Melanoma is a severe kind of skin cancer with a very high mortality rate. Although there are only 2% of all the skin cancer types, melanoma is responsible for 75% of deaths occurred due to skin cancer [1]. In USA only, about 87,110 new cases are reported every year out of which 9,730 patients lose their lives due to this lethal skin cancer [2]. Similarly, in 2016 a total of 6,800 fatalities due to melanoma were reported in Canada [3]. Usually, the exposed regions of skin to

sunlight are highly affected by melanoma e.g., face, legs, and arms. The borders and colors of melanoma moles are uneven and evolving which represent the severity level of the disease [4]. Many advanced techniques for the treatment of skin cancer are available including radiation therapy and immunotherapy. In clinical practice [5], these techniques are combined with surgery but still the survival rate of advanced stages of melanoma is quite low and is around 15%. On the other hand, the survival rate for the early stages of melanoma is around 95% [6]. In order to diagnose the medical

problems of melanoma lesions, dermatologists directly examine the damaged skin's uniformity, inconsistencies in the borders, and color changes [4]. Moreover, dermoscopy, a nontrauma skin imaging technique, is also very popular to assist dermatologists to examine the affected skin. The accuracy for identification of melanoma lesions through dermoscopy is higher than the traditional method of ABCD rule criteria [7]. This ABCD rule is designed by the American Society for skin lesions [8]. Nevertheless, the biopsy test is the only thing on which the performance is solely dependent. In the initial stages, the identification of melanoma greatly matters since in the initial stages the possibility of recovery is much higher than in the later stages. However, the manual identification of melanoma needs an expert dermatologist followed by a stage in which the decision is made to assess a subjective variation.

Numerous researchers have proposed to automate the analysis process and extend the knowledge that can identify lesions accurately and helps different healthcare systems which are based on the Internet of Things (IoT) [9–11]. There exist traditional techniques, e.g., Otsu and Stochastic, that can perform melanoma segmentation, but these thresholding techniques are not the end-to-end solution, and owing to artifacts, this might lead to under or over segmentation problems. Therefore, there is a need for automated systems to automatically diagnose skin lesions for the treatment of skin cancer patients. The lighting conditions and different orientations also make it a challenging task for automated systems to analyze them [12]. Some researchers highlighted these issues recently and observed that there is very low diagnostic accuracy due to the presence of these issues in clinical images [13].

Recently, deep-learning approaches are also utilized for the task of automated skin lesion segmentation to overcome the challenges with traditional methods. The performance of these deep-learning-based methods is exceptional in segmenting skin lesions as compared to the traditional dermatologists [1]. A lot of deep-learning-based segmentations are proposed in the existing research studies for skin lesions, but there is still space to enhance the algorithms in terms of both parameters and performance [14–17].

From this line of research, we proposed an efficient deep-learning model for end-to-end segmentation of melanoma lesions to overcome all the challenges which include intraclass variations and lighting conditions as well as other related issues. The proposed framework uses the UNet architecture as the base architecture for end-to-end segmentation of melanoma lesions, as it has a very strong capability in biomedical image segmentation [18]. More explicitly, it consists of a downsampling path, a bottleneck layer, and an upsampling path. The downsampling path consists of multidilated convolution blocks (MDC) and depthwise separable convolutions blocks (DSC) that empower the process of feature learning across the channels on the image. The parameters of convolution are dramatically reduced with these depthwise separable convolutions without compromising the performance. The generalization ability of the model is improved by these convolutions while

avoiding overfitting. Spatial and cross-channel correlations are also separated with the help of these convolutions. Moreover, there is the use of swish activations in the MDC block. The nonmonoatomic property of swish is very advantageous in deep-learning algorithms. All these characteristics make the proposed framework more reliable in segmenting melanoma lesions. The following points describe our contribution:

- (i) The proposed approach is capable of localizing melanoma lesions and multiple types of cancer in a single image by designing DSC blocks with multidilated features
- (ii) The proposed segmentation model accurately segments the lesions by overcoming the challenges presented in the ISBI2016 dataset
- (iii) We used skin refinement as a preprocessing step to eliminate artifacts from dermoscopic images.

The rest of the paper is organized in the following way: Section 2 thoroughly explains the review of current approaches. Section 3 explains our proposed methodology in detail. Section 4 explains the experimental details, results, and discussion. Lastly, Section 5 provides the conclusion of the paper.

## 2. Literature Review

The segmentation of melanoma lesions is a fundamental technique in designing the automated detection model of skin cancers. Since the segmentation of lesions plays an important role in the classification task of skin cancer [19–21]. Automated segmentation techniques are further split into traditional and deep-learning techniques, along with some advanced hybrid deep learning models. The following is a critical literature assessment of each kind of method in the segmentation of melanoma lesions.

The conventional techniques of melanoma lesions segmentation mostly involve iterative selection [22, 23], adaptive threshold [24], iteration merging of regions [25], and Otsu threshold [26]. Nevertheless, as a result of the existence of artifacts in dermoscopic images the effectiveness of thresholding-based techniques will be diminished [22, 26]. In [26], the accuracy of the proposed algorithm is acceptable but the images that were segmented have uneven borders as well as reduces the resolution of the images. In [27], the authors suggested a method to address the challenges that arise in [26]. Another collection of studies [25, 28] suggested a region merging technique to perform segmentation. In this method, the identical regions of the images are clustered together. To overcome the challenges of color, low contrast, and illustration, the region merging technique performs well. In [25], lesion segmentation is carried out by these identical regions having identical attributes. Overall, these approaches need a lot of manual parameter tuning, such as threshold values in thresholding-based segmentation, making them unsuitable for automated CAD systems.

Another research group [29–32] suggested deep-learning techniques for segmentation and achieved considerable outcomes as compared to the standard methods.

In [31], an FCRN, i.e., fully convolutional-residual-network was suggested to address the challenges of model overfitting in the task of melanoma segmentation. In [32], localization of lesions is accomplished by utilizing the region-based CNN followed by the machine learning fuzzy-clustering technique. In [29], a 19-layer CNN is designed to improve and enhance the results of melanoma segmentation. More specifically, in this study, Jaccard distance is utilized as a loss function. With the assistance of this loss function, the segmentation performance improves and also the problem of overfitting arises between normal and melanoma images. In [30], FRCN, i.e., full CNN was designed for segmentation of melanoma lesions. In order to segment, the lesion areas of different scales a segmentation model based on multiscale convolution is proposed in [33] which efficiently extracts the areas of lesions. A multistage segmentation model was proposed in [34] to perform the end-to-end segmentation of skin lesions. They also combined and integrates the context information with their model. The boundary of lesion segmentation is further improved in [35]. They combined the mixed feature inputs and proposed a multibranch fusion network and performed an immense set of experiments to evaluate their model. In [36], a new method for automatic segmentation of skin lesions is designed which was capable of learning more powerful and distinguishable features. This model used cross-net-based aggregation. In [37], to segment lesions and lessen the impact of artifacts, a hybrid technique was suggested by integrating the convolutional and recurrent neural networks. Nevertheless, a two-stage object detection model such as RCNN produces about 2 thousand patches per image for lesion identification. Due to this reason, melanoma localization becomes computationally expensive in these approaches. Furthermore, while all of these deep-learning algorithms for lesion segmentation produce outstanding results, there is still a gap for improvement in terms of model performance.

In addition, to acquire more information features from dermoscopy images, some hybrid models are also designed such as in [38] for bilinear merging, they used ResNet and VGG to extract high-level features and trained their algorithm using SVM classifiers. They achieved the best accuracy results on several test sets. In order to cope with the intraclass inconsistency of lesions, a multi-convolution neural network is proposed in [39]. This model was combined with an adaptive sample strategy of learning. This technique also deals with related noise interference. In [40], encoded output features are converted into Fisher Vectors by using the weights of the pretrained model which is a deep residual network. They also used trained SVM to achieve the recognition task and have achieved a significant performance on a test set of classification challenges of ISBI2016. However, their approach was not an end-to-end solution and the overall architecture of the model was very complex. The advantages of hybrid approaches include improved performance and broader feature acquisition; nevertheless, the computational complexity of hybrid deep-learning models is high, making them slow.

### 3. Methodology

The detail of our proposed framework is presented in Figure 1. In this research, we have utilized three different datasets. The instances in the dataset undergo some pre-processing stages for improved quality images to remove artifacts like hair, bubbles, and other patches. This is followed by steps to localize the melanoma lesions.

*3.1. Preprocessing.* Before giving the input images to the deep-learning model, all the images are preprocessed to remove noises from them. This step is necessary for very precise segmentation. Most commonly used image pre-processing techniques involve image smoothing, resizing, identification of ROIs, and denoising of images. For the elimination of artifacts from dermoscopic images, Gaussian smoothing is the most effective technique. In the suggested method, we have performed the dilation followed by erosion also referred to as morphological closing. Later on, in the next stage, we performed the sharpening operations over the images to further enhance the quality of the images. Some sample images before and after preprocessing are depicted in Figure 2.

*3.2. Data Augmentation.* Usually, the publicly accessible training images for all categories are not dispersed evenly, resulting in the class imbalance issue [41]. In the suggested method, we increase the total number of samples in the train set by employing different types of augmentation such as flipping, cropping, and rotating. Table 1 lists the different types of augmentation and their values used to augment the samples. More specifically, 15 additional images are sampled from a particular dermoscopic image by using the augmentation types given in Table 1. The main rationale to use this phase in our strategy is to reduce overfitting problems and improve the model's predictive performance.

*3.3. Proposed Architecture.* Our proposed framework consists of three major parts which include the downsampling path to down sample an image by extracting the features which represent what is present in an image followed by the bottleneck and upsampling path to upsample an image to get the localization of the required lesion in an image as shown in Figure 3(a). The complete architecture of each part is described below:

*3.3.1. Downsampling Path.* The downsampling path of the model consists of a multidilated convolution (MDC) block and depthwise separable [30] convolution block (DSC) to encode features of melanoma lesions followed by max-pool operations of size  $2 \times 2$  to reduce the spatial dimensions of the images as shown in Figure 3(a). The architecture of the MDC and DSC block are given in Figure 3(b). The feature extraction part starts from the regular convolution of size  $1 \times 1$  and  $3 \times 3$  max-pool on an input image of size  $256 \times 256 \times 3$  followed by ReLu [33] activation functions. Besides this, input is also given to the DSC block as shown in



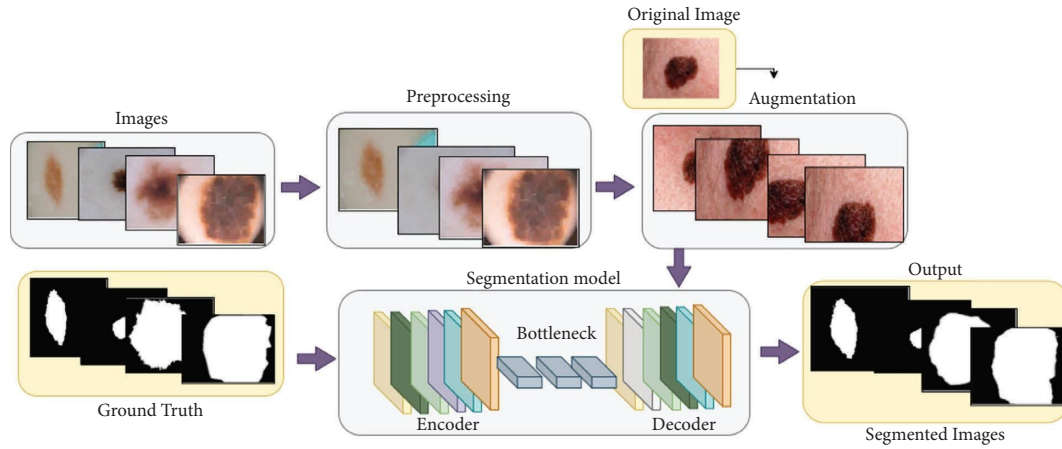


FIGURE 1: A schematic overview of the proposed methodology.

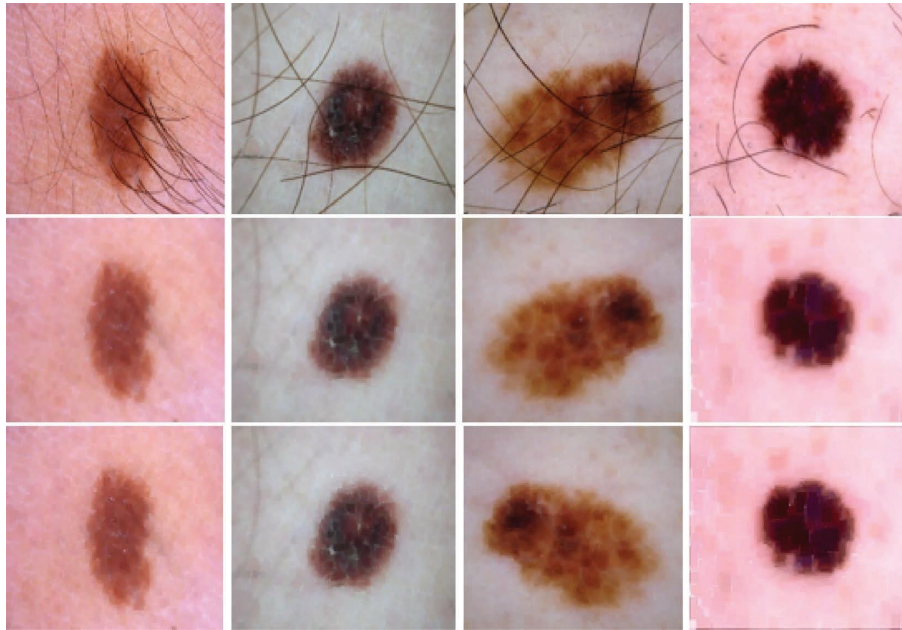


FIGURE 2: Image enhancement on ISBI2016 dataset; row 1 depicts the original dataset images, row 2 depicts the images after closing morphological operation, and row 3 depicts results after sharpening.

TABLE 1: Data augmentation types and their parameters.

Sl. no.	Augmentation types	Parameters
1	Rotate	90°, 180°, 270°
2	Crop from right	45°, 60°, 90°
3	Crop from left	45°, 60°, 90°
4	Crop from top	45°, 60°, 90°
5	Crop from bottom	45°, 60°, 90°
6	Flipping	Left right
7	Shifting	Shifted by (25, 25) pixels

Figure 3(b). In the DSC block, the depthwise separable convolution of sizes  $1 \times 1$  and  $3 \times 3$  is performed on every channel of an input image independently. Afterwards, a  $1 \times 1$  window is utilized as pointwise convolution to project to a new channel space after a channel is computed

by depthwise convolution as shown in Figure 4. The depthwise separable convolutions are not like spatial separable convolutions which are also referred as “separable convolutions” in the community of image processing [42]. The mathematical formulation is given below:

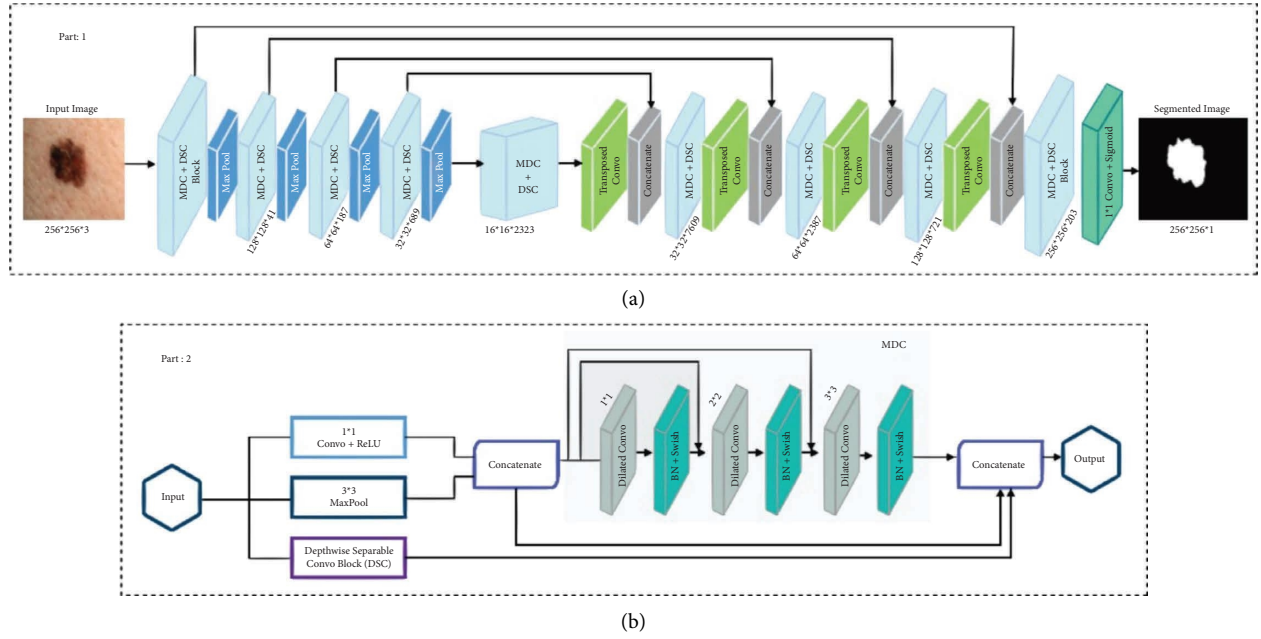


FIGURE 3: Architecture of proposed segmentation model.

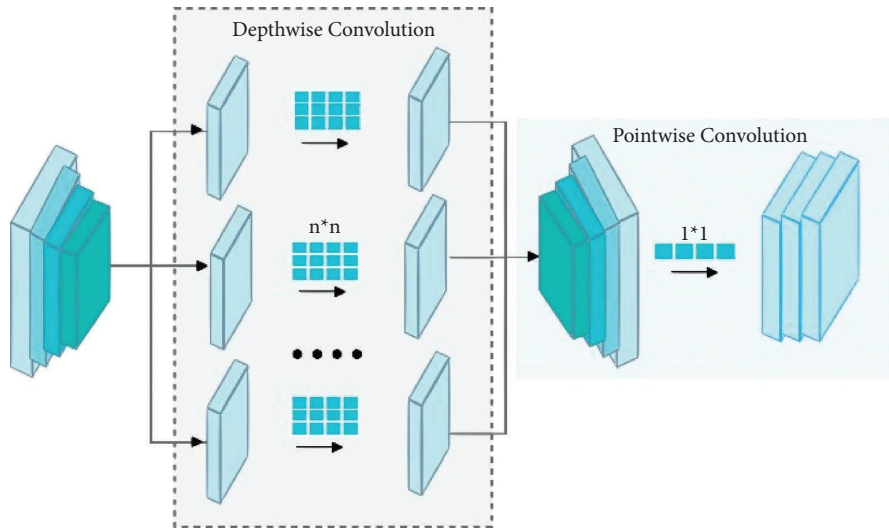


FIGURE 4: Depthwise separable convolutions.

$$\begin{aligned}
 \text{Conv} &= (W, y)_{(i,j)} = \sum_{k,l,m}^{K,L,M} W_{(k,l,m)} \cdot y_{(i+k,j+l,m)}, \\
 \text{PointWiseConv} &= (W, y)_{(i,j)} = \sum_m^M W_m \cdot y_{(i,j,m)}, \\
 \text{DepthWiseConv} &= (W, y)_{(i,j)} = \sum_{k,l}^{K,L} W_{(k,l)} \odot y_{(i+k,j+l)}, \\
 \text{SeparableConv} &= (W_p, W_{d,y})_{(i,j)} = \text{PointWiseConv}_{(i,j)}(W_p, \text{DepthWiseConv}_{(i,j)}(W_{d,y})).
 \end{aligned}
 \tag{1}$$

In the above equations,  $\odot$  shows the elementwise product. The benefit of depthwise separable convolutions over traditional convolutions is the total number of parameters [43]. For this, consider a standard convolution with a feature map  $\mathbf{F}$  and suppose that value of stride and padding is one. This can be computed as the following equation:

$$\mathbf{F}_{k,l,m} = \sum_{i,j,m} \mathbf{K}_{i,j,m,n} \cdot \mathbf{I}_{k+i-1,l+j-1,m}. \quad (2)$$

For these standard convolutions, the total number of parameters and computational cost can be calculated as follows:

$$\mathbf{k} \times \mathbf{k} \times \mathbf{M} \times \mathbf{N} \text{ and } \mathbf{k} \times \mathbf{k} \times \mathbf{M} \times \mathbf{N} \times \mathbf{H} \times \mathbf{W}, \quad (3)$$

where the input image or input feature maps are represented by  $\mathbf{I}$ , while  $\mathbf{k}$  denotes the kernel of convolution with size  $\mathbf{k} \times \mathbf{k}$ . The  $\mathbf{M}$  and  $\mathbf{N}$  denote the number of input and output channels while the height and width of input feature maps or input images are denoted by  $\mathbf{H}$  and  $\mathbf{W}$ , respectively. Furthermore, for depthwise separable convolutions which is a combination of depthwise and pointwise convolutions, the output feature maps are calculated as follows:

$$\mathbf{F}_{k,l,n} = \sum_{i,j} \mathbf{K}_{i,j,m} \cdot \mathbf{I}_{k+i-1,l+j-1,m}. \quad (4)$$

Similarly, for these depthwise separable convolutions, the total number of parameters and computational cost is calculated as follows:

$$\mathbf{k} \times \mathbf{k} \times \mathbf{M} + \mathbf{M} \times \mathbf{N} \text{ and } \mathbf{k} \times \mathbf{k} \times \mathbf{M} \times \mathbf{H} \times \mathbf{W} + \mathbf{M} \times \mathbf{N} \times \mathbf{H} \times \mathbf{W}. \quad (5)$$

Now, in order to compare the parameters of both types of convolutions, we obtained the following equation:

$$\frac{\mathbf{k} \times \mathbf{k} \times \mathbf{M} + \mathbf{M} \times \mathbf{N}}{\mathbf{k} \times \mathbf{k} \times \mathbf{M} \times \mathbf{N}} = \frac{1}{\mathbf{N}} + \frac{1}{\mathbf{k}^2}. \quad (6)$$

It can be shown and seen that the number of parameters is about 8 to 9 times less in depthwise separable convolutions than in standard convolutions. Hence, it is observed that we improved the network without an extensive increase in the number of parameters of the network and also empowered the network to learn deep dilated features which in turn gives more contextual information. Moreover, the output of regular convolutions and max-pool are concatenated and given as input to the first dilated convolution in the MDC block as shown in Figure 3(b). Similarly, the input of second and third dilated convolutions in the MDC block is the output of regular convolutions, max-pool, and the result of

previously dilated convolution. Furthermore, in the MDC block, three convolution operations utilizing the dilated filters of size  $1 \times 1$ ,  $2 \times 2$ , and  $3 \times 3$ , respectively, are used. The convolutions which use the dilated filters are also called dilated or atrous convolutions. For these, a dilated filter  $\mathbf{w}$  also called kernel is convolved over the input signal, and for each location,  $\mathbf{i}$  is the output, and  $\mathbf{y}$  is computed by equation (7), ([44])

$$\mathbf{y}[\mathbf{i}] = \sum_{\mathbf{k}} \mathbf{x}[\mathbf{i} + \mathbf{r} \cdot \mathbf{k}] \mathbf{w}[\mathbf{k}]. \quad (7)$$

In equation (7) the  $r$  is representing the value of the stride by which the input signal is sampled which is a similar operation to convolve over any input signal  $x$  with the help of filters  $w$  that are upsampled by inserting  $r - 1$  zero along each spatial dimension that are consecutive. These are very helpful as a large receptive field of view is enhanced by dilated convolutions of the given input image. After each dilated convolution in the MDC block, there is the use of batch normalization [31] and swish activations [32] as shown in Figure 3(b). The use of batch normalization [31] fastens the training process and prevents the model from overfitting. A dropout layer of rate 0.05 is also added after every max-pool operation. Furthermore, the swish activations are defined as [45]

$$f(x) = x \cdot \sigma(x). \quad (8)$$

In equation (8), the  $\sigma(x) = (1 + \exp(-x))^{-1}$  represents the sigmoid function. This activation function is bounded below and unbounded above. The properties of swish activation include that it is smooth and the property of non-monotonicity which distinguishes it from other activation functions. The derivative of the swish is given below in equation (9) [45]

$$\begin{aligned} \mathbf{f}(\mathbf{x})' &= \sigma(\mathbf{x}) + \mathbf{x} \cdot \sigma(\mathbf{x})(1 - \sigma(\mathbf{x})) \\ &= \sigma(\mathbf{x}) + \mathbf{x} \cdot \sigma(\mathbf{x}) - \mathbf{x} \cdot \sigma(\mathbf{x})^2 \\ &= \mathbf{x} \cdot \sigma(\mathbf{x}) + \sigma(\mathbf{x})(1 - \mathbf{x} \cdot \sigma(\mathbf{x})) \\ &= \mathbf{f}(\mathbf{x}) + \sigma(\mathbf{x})(1 - \mathbf{f}(\mathbf{x})). \end{aligned} \quad (9)$$

Moreover, the output of MDC blocks is concatenated to depthwise convolution blocks, and the result of regular convolutions and max-pool is shown in Figure 3(b). The number of filters set for each of our convolution blocks is 16, 32, 64, and 128, respectively. Moreover, the starting weights for regular convolution and convolutions in MDC blocks are initialized with ‘‘He normal’’ weight initialization which is defined as [46, 47]

$$\mathbf{W} \sim \mathbf{G}\left(\mathbf{0}, \sqrt{\frac{2}{\mathbf{n}}}\right) \quad (10)$$

Or  $\mathbf{W}[\mathbf{i}] = \text{RandomUniform}(\text{low} = -\text{limit}, \text{high} = \text{limit}, \text{size} = (\mathbf{F}_{\text{in}}, \mathbf{F}_{\text{out}}))$ .

In the above equation (10),  $\mathbf{G}$  is just a random number with Gaussian probability distribution while the total number of inputs coming towards a particular neuron is represented by  $\mathbf{n}$ . Furthermore,  $\sqrt{2/\mathbf{n}}$  is used to calculate the standard deviation while the 0 represents the mean. In addition,  $\mathbf{F}_{in}$  and  $\mathbf{F}_{out}$  are the number of inputs and outputs to the layer, respectively. Similarly, the weights of depthwise separable convolutions are initialized with the Glorot weight initialization method which is also called Xavier initialization. The main objective of the downsampling path is to extract features that describe the semantics of the image with loss of spatial and localization information.

**3.3.2. Bottleneck Path.** The bottleneck path of the proposed framework consists of  $1 \times 1$  and  $3 \times 3$  convolution followed by depthwise separable convolution block (DSC) and MDA blocks as shown in Figure 3(a). The resulting feature maps of the last max-pool operation on an input image in the downsampling path are given as inputs to the bottleneck path which yields output feature maps of dimension  $16 \times 16 \times 2323$ . These resulting feature maps are then given as input to the very the first layer of the upsampling path to localize the melanoma lesion.

**3.3.3. Upsampling Path.** The upsampling path of the model consists of transposed convolution with kernel sizes of  $3 \times 3$  with a stride of  $2 \times 2$  followed by the operation of concatenation to corresponding convolution blocks of downsampling path as shown in Figure 3(a) to combine the context and localization information to segment out the melanoma lesions. Transposed convolutions are the reverse processes of convolution, and it is more robust than simple upsampling as it fills up the details with proper learning. These are also called fractionally stride convolutions. Moreover, the concatenation operations between upsampling and downsampling path at the appropriate position help to restore the localization information that is lost during downsampling an image. So more specifically, the input from the bottleneck layer is first given as an input to the first transposed convolution layer. Then, by means of skip connections, the output generated from this layer is concatenated to the last MDC and DSC blocks downsampling path. Moreover, this process is repeated three more times. In the end, the output of the last MDC and DSC blocks in upsampling path is passed through  $1 \times 1$  convolution followed by sigmoid activation to get the required segmented image of the lesion.

## 4. Experiments, Results, and Discussion

In this section, we discuss the datasets used for experimentation purposes and evaluation metrics used to evaluate the model as well as results of the model. In addition, the proposed model is designed in the Keras framework available in Python, and simulations are run on Google Colab with 12 GB RAM and NVIDIA Tesla K80 GPU. The hyperparameters of the model include the weight initialization, weight optimizer, learning rate, and epochs which are set to Xavier, Adam, 0.001, and 150, respectively.

**4.1. Datasets.** To assess the universality of our proposed model, we evaluated it on three distinct datasets, i.e., DermIS, DermQuest, and ISBI2016. All the datasets contain skin lesion images in RGB format. More explicitly, the DermQuest contains 152 melanoma images while 122 images belong to the nevus class. Similarly, in DermIS, the total number of melanoma class images is 43 while the nevus class has a total of 26 images. The DermQuest and DermIS datasets contain a limited number of images, so augmentation is applied to the training set. Moreover, the dataset ISBI2016 comprised 900 melanoma images in the train set and 379 images in the test set. The train and test division of images are already provided by the dataset publisher. For a fair comparison, we utilize the same train and test sets.

**4.2. Performance Evaluation Metrics.** To examine the performance of the model, we utilized different evaluation metrics [48–51] including dice score, specificity, sensitivity, and Jaccard score. The following equations (11)–(15) are used to compute these metrics

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \quad (11)$$

$$\text{Dice score} = \frac{2 \times \text{TP}}{2 \times \text{TP} + \text{FP} + \text{FN}}, \quad (12)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}, \quad (13)$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (14)$$

$$\text{Jaccard score} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}}, \quad (15)$$

where TP denotes the true positives, FP denotes the false positives, TN denotes the true negatives, and FN denotes the false negatives.

**4.3. Results of DermIS Dataset.** In the first step, we evaluate the proposed model on the DermIS datasets containing melanoma and nevus class images along with their mask images. As previously stated, artifacts like hair, air bubbles, and other noises can be seen in the images of the DermIS dataset. The existence of these types of artifacts will influence performance accuracy. To address this problem, we have performed the preprocessing on images that are discussed in Section 3.1. In addition, we have also performed the data augmentation described in Section 3.2 to increase the number of training samples since DermIS has a very limited number of images. This is done to expand the number of instances since a minimal amount of training data leads to overfitting issues. In Figure 5, the results of augmentation are depicted. The proposed model takes the dermoscopic images along with their ground truth images as input and outputs the segmented images. The results of melanoma segmentation are depicted in Figure 6 along with their actual ground truth images and contour images.



FIGURE 5: Results of augmentation on DermQuest dataset.

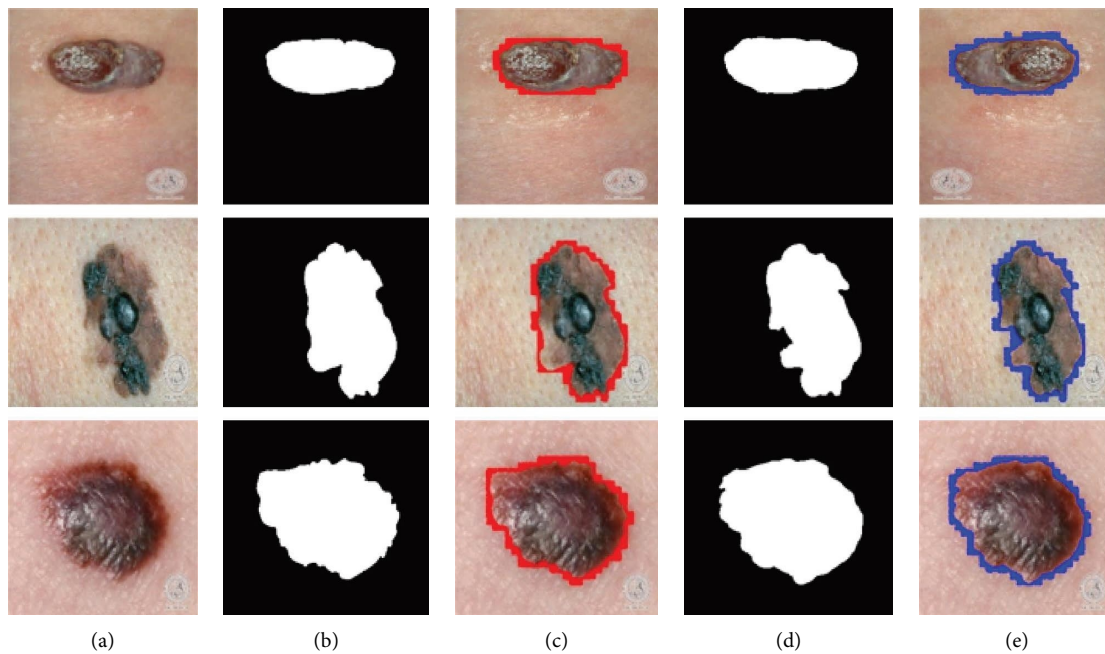


FIGURE 6: Results of melanoma segmentation on DermIS dataset.

Column (A) in Figure 6 shows the original images that were preprocessed. Column (B) shows the actual ground truth images. Following on, column (C) shows the contour images of actual ground truth images. The contour is shown by the red borders in column (C). Column (D) depicts the output of the segmentation model in form of segmented images while column (E) shows the output images with contours. The effectiveness of the proposed method on this database was assessed utilizing previously defined metrics. As shown in Table 2, the Dice score achieved for this dataset is 97% which shows the robustness of our model performance in localizing skin lesions. The accuracy and Jaccard indexes are 97% and 94% while sensitivity and specificity are 93%.

*4.4. Result of DermQuest Dataset.* In the second step, we evaluated the performance of the proposed model on the DermQuest dataset. All of the trials on this data, like the DermIS dataset, make use of melanoma images and associated ground truth images. More specifically, we first perform the preprocessing step over the images to eliminate the noises in the form of artifacts. The number of images in this dataset is also less in number; hence, we also perform the data augmentation on this dataset. The results of melanoma segmentation for the DermQuest dataset are depicted in Figure 7 along with their actual ground truth images and contour images. Column (A) in Figure 6 shows the original images, column (B) shows the actual ground truth image, and column (C) shows the contour images of actual ground

TABLE 2: Results of skin lesion segmentation in ISBI2016 challenge.

Technique	Accuracy	Dice score	Jaccard score	Sensitivity	Specificity
ExB	0.95	0.91	0.84	0.91	0.965
CUMED	0.94	0.897	0.829	0.911	0.957
Mahmudur	0.952	0.895	0.822	0.88	0.969
SFU-mial	0.944	0.885	0.811	0.915	0.955
TMU team	0.946	0.888	0.81	0.832	0.987
UiT seg	0.939	0.881	0.806	0.863	0.974
IHPC-CS	0.938	0.879	0.799	0.91	0.941
UNIST	0.94	0.867	0.797	0.876	0.954
JoseLuis	0.934	0.869	0.791	0.87	0.978
Marco Romelli	0.936	0.864	0.786	0.883	0.962
Proposed	0.95%	0.90% (None)	0.82% (None)	0.92%	0.90%
	0.95%	0.947% (mi)	0.90% (mi)	0.92%	0.90%
	0.95%	0.92% (ma)	0.86% (ma)	0.92%	0.90%

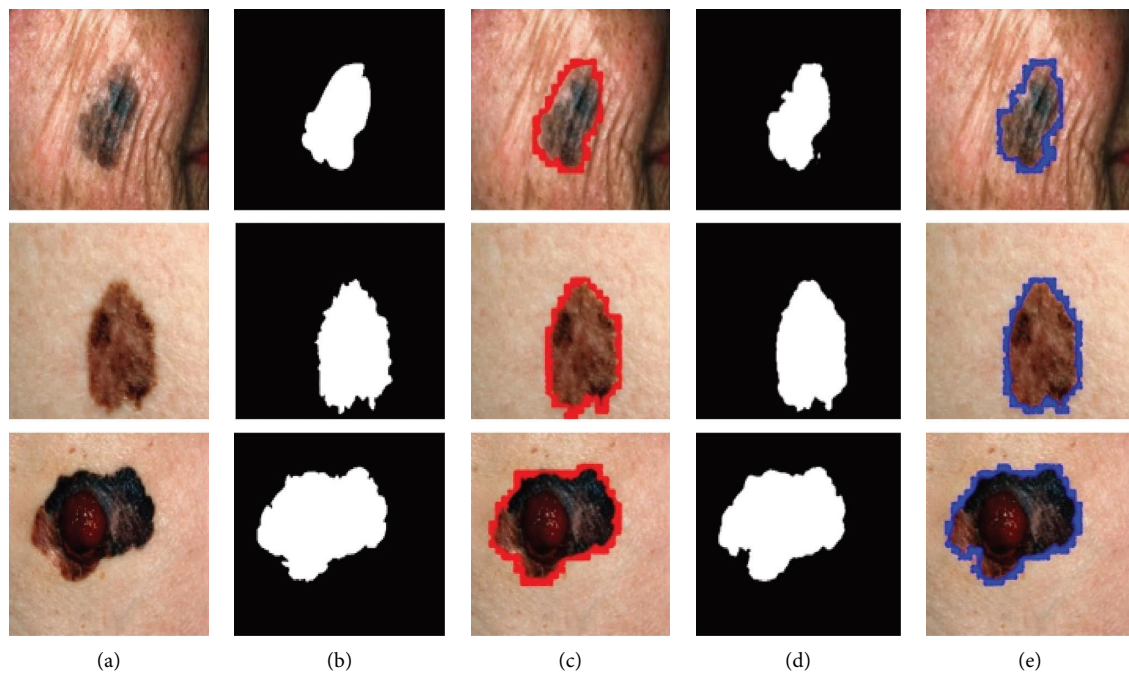


FIGURE 7: Results of melanoma segmentation DermQuest dataset.

truth images. Moreover, column (D) depicts the output of the segmentation model in form of segmented images while column (E) shows the output images with contours. For this dataset, we have achieved the highest Dice score, accuracy, and Jaccard score in comparison with the DermIS dataset. The proposed model achieved the Dice score of 97% and the Jaccard score of 96% in localizing the melanoma lesions. Moreover, the accuracy, sensitivity, and specificity attained for this dataset are 98%, 90%, and 95%, respectively.

**4.5. Results of ISBI2016 Dataset.** The suggested framework's efficacy was also examined using benchmark datasets namely ISBI 2016 by "International Symposium on biomedical images (ISBI) in the challenge of skin lesion analysis towards melanoma detection" [52]. For the challenge of segmentation, this database comprises a total of 1,279 images out of

which 900 images belong to the train set while the remaining 379 images belong to the test set. All dermoscopic images in this dataset, like those in DermIS and DermQuest, go through the preprocessing stage. The total number of training images in this dataset is sufficient for training purposes; hence, there is no data augmentation is applied to this dataset. Figure 8 shows the segmentation results of the proposed algorithm on the ISBI2016 dataset. In Figure 8, column (A) shows the original test images with their ground masks shown in column (B). The test images with contour around the boundary are shown in column (C). The predicted mask and output with contour are shown in columns (D) and (E) of Figure 8, respectively. In the test set of this dataset, there are more challenging images. As shown in row 1 of Figure 8, the lesion area of the first image has very similar to normal skin but still, it can be accurately segmented by a model as shown in row 1 column (D) of

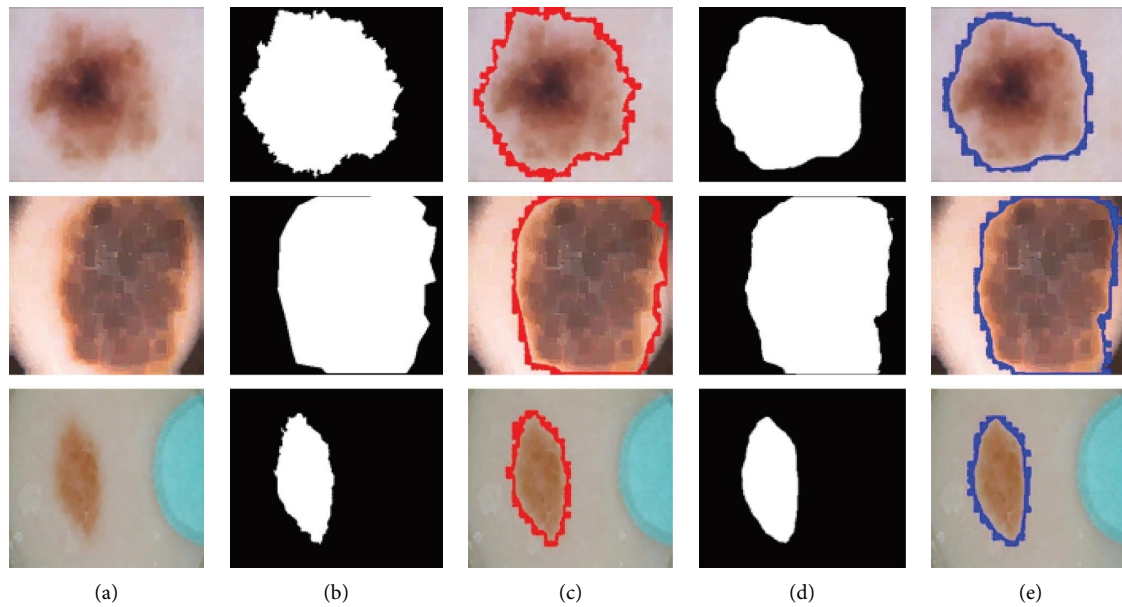


FIGURE 8: Sample melanoma segmentation results of ISIC2016 dataset from the skin with the respective masks and contour images.

Figure 8. The boundaries of lesions are still more distilled and smooth. The evaluation scores achieved by our proposed model on this dataset include a Dice score which is 94.7%, a Jaccard score of 90%, and an accuracy of 95%, respectively. Moreover, the sensitivity and specificity achieved for this dataset are 92% and 90%, respectively.

We also compared our results with challenge winners of ISBI2016. In this challenge, almost 28 groups provide their results, as listed in Table 2. This ISBI ranked the competition participants based on their best average Jaccard score. Due to the precise segmentation of deep-learning models, it is observed from Table 2 that most of the participants in the competition employ deep-learning techniques. For instance, AlexNet, VGG16, and ResNet-based pretrained models are utilized to approximate the edges and boundaries of lesions.

It is evident from Table 2 that the proposed algorithm attained the highest results among challenge winners. The comparison with all challenge winners and the proposed framework is given in Table 2 and is graphically presented in Figure 9. In terms of the Jaccard score, the proposed model has a very remarkable performance over the top two participants. The Dice score of the proposed model is also improved among all challenge winners. Moreover, the scores of each test set image in the ISBI2016 dataset are shown in Figure 10. It is observed from Figure 9, that most of the test samples achieved greater than 80% Dice, Jaccard, and accuracy scores. There are only a few samples in which the Jaccard score falls below 50%. Moreover, to consider the effect of class unbalancing, we calculate the Dice and Jaccard score in three different ways. First, we consider no averaging method and calculate the scores; in the second way we consider the average method of “micro” (mi) which globally calculates the FP, FN, and TP without favoring any class. Similarly, in the third way, we use the average method of “macro” (ma) in which we calculate the scores separately for

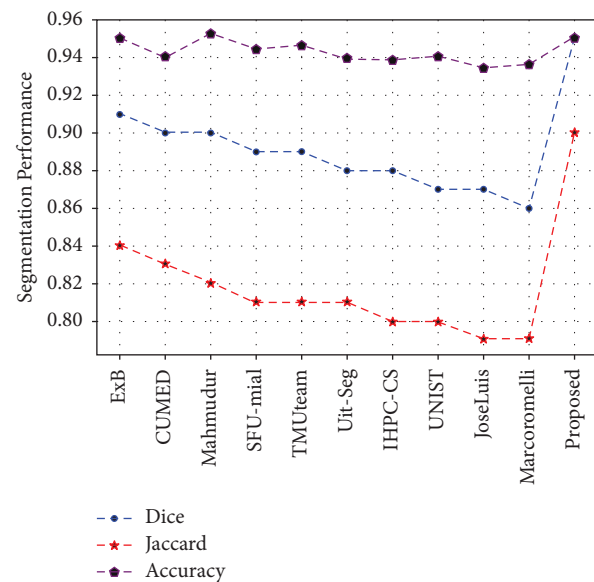


FIGURE 9: Comparison in terms of accuracy, Jaccard, and Dice scores with challenge winners.

both background and foreground classes. It is observed from the results that our proposed framework significantly addresses the challenges of segmentation in skin lesions which includes intraclass differences and visual similarity of lesion features with normal skin.

Furthermore, the training graphs of accuracy and loss of the model for all three datasets are also shown in Figure 11. In general, the accuracy of the model is used to determine the total number of correct predictions. The higher value of accuracy shows the better capability and performance of the model. The graphical representation of accuracy is shown in Figure 11, and it is observed that during training the model

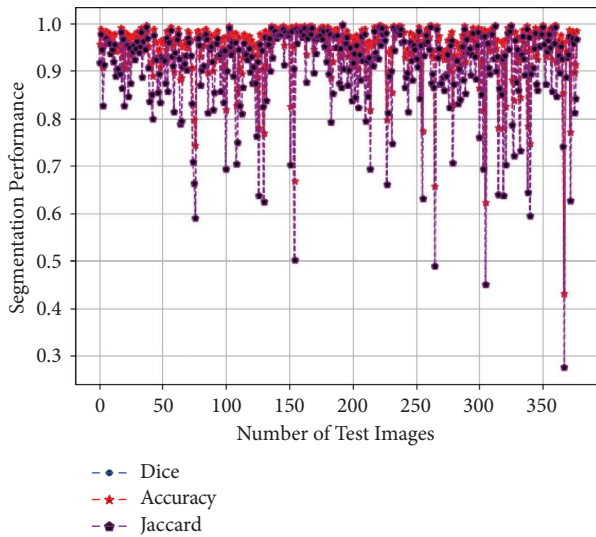


FIGURE 10: Segmentation performance of each test image in the ISBI2016 dataset.

achieves an accuracy greater than 90%. Similarly, model loss values during training of all three datasets are also plotted. The predictions of the model are more accurate if the loss of the model is near to zero. It is observed that the loss values of the proposed model on all three datasets are near zero. The x-axis of Figure 11 shows the total number of epochs while the y-axis shows the accuracy and loss values epoch by epoch of the proposed model. Moreover, during the training of the deep-learning model, when an input image passes through successive layers of architecture; then, each layer gives output in the form of feature maps of different dimensions. These feature maps indicate how your model encodes and learns the features of images layer by layer. Usually, in the starting layers, the model extracts low-level features while subsequently more high-level features are extracted. The activation maps of some intermediate layers of the proposed algorithm are also shown in Figure 12.

Figure 12 illustrates that lesion areas are more focused on the proposed model. This indicates that the model learns more effective and discriminative features of lesion areas in the given image.

#### 4.6. Comparative Analysis with State-of-the-Art Approaches.

We have compared the performance accuracy of our proposed framework with other state-of-the-art approaches. It is noticeable from Table 3 that recent approaches use many deep-learning approaches to automatically segment melanoma lesions. Bozorgtabar et al. [53] proposed an unsupervised method for skin lesion segmentation. In this work, the information about the context of the image is exploited at the superpixel level. They achieved Dice and Jaccard scores of 0.86% and 0.66%, respectively. Similarly, Yaun et al. [29] proposed a 19-layer deep convolutional network for automatic segmentation of skin lesions. In their work, the proposed model is trained with a loss function of Jaccard distance and achieved Dice and Jaccard scores of 91% and 84%, respectively, which is very much better.

Furthermore, Li et al. [43] proposed a dense convolutional neural network based on residual learning for skin lesion segmentation. They achieved a Dice score of 93% with an 87% Jaccard score. Rashid et al. [40] proposed a two-stage method and utilized the approach of object detection algorithms named single shot detector (SSD) for localization of melanoma lesion followed by a second stage in which level set algorithm is used to segment the melanoma lesion. The Jaccard and Dice scores achieved by their approach are 90% and 82%, respectively. Moreover, Tang et al. [34] proposed a new novel multistage UNet-based model combined with context information fusion structure (CIFS) for melanoma segmentation and achieved an appropriate improvement in the Jaccard score. In comparison with all the previous approaches, our model outperforms especially in terms of Jaccard analysis. Wei et al. [45] proposed an ensemble lightweight neural network for melanoma segmentation and achieved a significant and excellent performance in Dice and Jaccard scores which are 96% and 92%, respectively. The main reason for having efficient performance results is the end-to-end automatic segmentation of melanoma lesions by employing the use of (DSC) blocks with multidilated filters which enlarges the receptive field and view of filters. Moreover, the nonmonoatomic property of swish activation makes the training smooth. Furthermore, in our approach, we applied a preprocessing technique on images that removes the artifacts in data that hinder the accurate segmentation of melanoma.

Table 3 represents the comparison between the existing techniques and the proposed framework. From Table 3, it is observed that there is significant improvement found in terms of Jaccard and Dice scores, especially in ISBI2016, which contains 379 challenging test images.

**4.7. Discussion.** Melanoma lesion segmentation remains one of the most difficult tasks in dermoscopy image analysis. Traditional segmentation methods such as Otsu and thresholding perform well but fails when artifacts and noises are observed in the images. In addition, they also require manual tuning of parameters such as threshold values. These manual settings limit their use in automated CAD systems. More explicitly, in CAD systems end-to-end solutions are preferable. Hence, in this research study, we proposed a deep-learning-based segmentation model to automatically segment the lesion from given dermoscopic images. The proposed model first encodes the dermoscopic images to extract the features of melanoma lesions using a DSC block in which depthwise separable convolutions are applied channelwise and has a smaller number of weights in comparison with the conventional convolutions. Following the activation function, swish is applied to achieve the nonlinearity on the resulting feature maps. In subsequent steps, the bottleneck layers are inserted followed by an upsampling path called a decoder to generate the segmented image containing the lesions. The proposed model performs well since it avoids the problems of overfitting by using convolution layers with fewer parameters using DSC blocks as well as by disentangling spatial and cross-channel



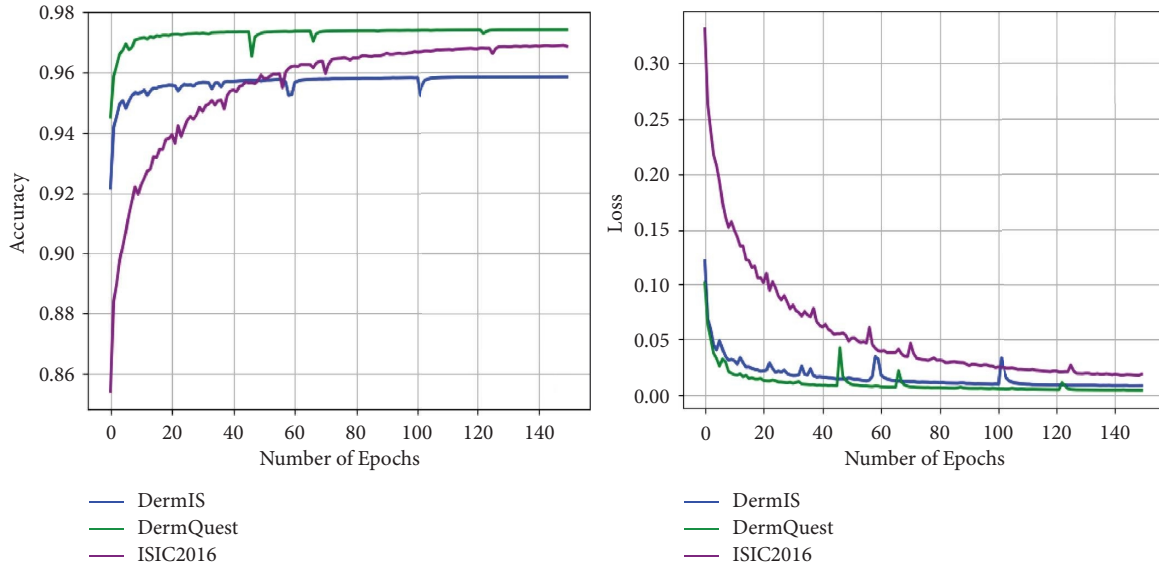


FIGURE 11: Loss and accuracy graphs of each dataset during training.

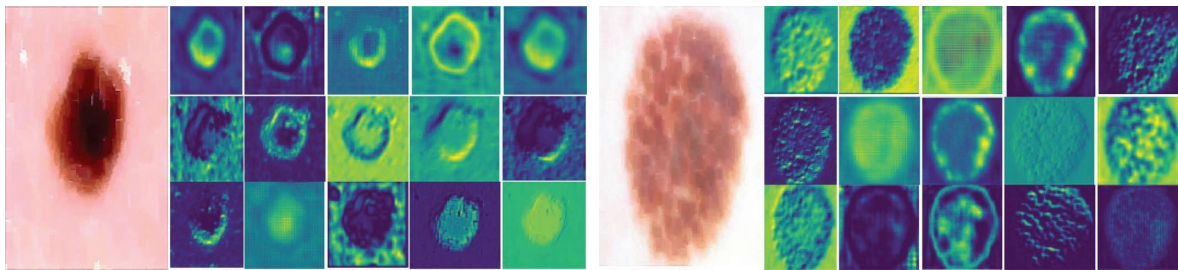


FIGURE 12: Results of channel activation of intermediate layers of the model.

TABLE 3: Comparison with state-of-the-art approaches.

Techniques	Accuracy	Dice score	Jaccard score	Specificity	Sensitivity
Rashid et al. [40]	0.90	0.901	0.82	0.98	0.83
Yaun et al. [29]	0.955	0.912	0.847	0.966	0.918
Li et al. [43]	0.959	0.931	0.870	0.96	0.95
Wei et al. [44]	0.962	0.923	0.867	0.974	0.934
Tang et al. [34]	0.95	0.91	0.85	0.96	0.92
Bozorgtabar et al. [53]	—	0.86	0.67	—	—
DermQuest	0.98	0.97	0.96	0.95	0.90
DermIS	0.972	0.97	0.94	0.93	0.93
ISBI2016	0.95	0.947	0.90	0.90	0.92

correlations. The results presented in Table 2 provide accuracy, Dice score, Jaccard score, sensitivity, and specificity of the proposed model in comparison with challenge winners of the ISBI2016 dataset. Similarly, Table 3 provides a comparison with different research studies. The proposed method's strength is that it accurately segments out lesions from dermoscopic images of not only melanomic type cancer but also nevus type cancer whose images are available in DermQuest and DermIS datasets. This indicates the generalizability of the proposed method in terms of segmenting different types of lesions. In addition, the proposed

method is less complex in comparison with the hybrid models that are large in terms of parameters. This is due to the adoption of DCS blocks in which depthwise separable convolutions are used to extract features with fewer number weights. However, one potential limitation of the method is that the model training is done from scratch, which takes long time for optimal convergence; thus, what if the encoder is set to pretrained weights? This would be an excellent future direction for this work. Furthermore, more challenging ISBI datasets on skin cancer should be utilized to investigate the performance.

## 5. Conclusion

Melanoma lesion segmentation is a very challenging task in the medical imaging domain since the normal and affected regions have the same appearance, and usually, the presence of artifacts and other noises in data decreases the segmentation performance. To address this challenge, different traditional segmentation methods are suggested by various researchers; however, these methods are not suitable for automated CAD systems due to many manual parametric steps. Therefore, we proposed a deep-learning-based segmentation model for automated segmentation of melanoma lesions from dermoscopic images. The suggested model employs the depthwise separable convolution blocks (DSC) which can learn the features from each space of an image. Moreover, multidilated filters broaden the view of kernels or filters and capture the information with large receptive fields. The use of swish activation proved to be very beneficial due to its nonmonoatomic behavior. The experimentation has been done on three different datasets including DermIS, DermQuest, and ISBI2016 datasets. The Dice and Jaccard scores for DermIS are 97% and 94%, for DermQuest are 97% and 96%, and for ISIC2016 are 94.7% and 90%, respectively. Future work will entail in improving the segmenting model by adding the attention modules such as CBAM and expanding the number of samples in training data in terms of challenging images.

## Data Availability

The data used to support the findings of this study are publicly available.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea Government (MSIT) (No. 2021R1G1A1095460) and also by the Chung-Ang University, Research Scholarship Grants in 2021.

## References

- [1] A. Esteva, B. Kuprel, R. A. Novoa et al., "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, 2017.
- [2] R. L. Siegel, K. D. Miller, S. A. Fedewa et al., "Colorectal cancer statistics, 2017," *CA: A Cancer Journal for Clinicians*, vol. 67, no. 3, pp. 177–193, 2017.
- [3] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2016," *CA: A Cancer Journal for Clinicians*, vol. 66, no. 1, pp. 7–30, 2016.
- [4] F. Nachbar, W. Stolz, T. Merkle et al., "The ABCD rule of dermoscopy: high prospective value in the diagnosis of doubtful melanocytic skin lesions," *Journal of the American Academy of Dermatology*, vol. 30, no. 4, pp. 551–559, 1994.
- [5] C. A. Barker and M. A. Postow, "Combinations of radiation therapy and immunotherapy for melanoma: a review of clinical outcomes," *International Journal of Radiation Oncology, Biology, Physics*, vol. 88, no. 5, pp. 986–997, 2014.
- [6] C. M. Balch, J. E. Gershenwald, S.-J. Soong et al., "Final version of 2009 AJCC melanoma staging and classification," *Journal of Clinical Oncology*, vol. 27, no. 36, pp. 6199–6206, 2009.
- [7] M. E. Vestergaard, P. Macaskill, P. E. Holt, and S. W. Menzies, "Dermoscopy compared with naked eye examination for the diagnosis of primary melanoma: a meta-analysis of studies performed in a clinical setting," *British Journal of Dermatology*, vol. 159, no. 3, pp. 669–676, 2008.
- [8] N. Razmjoooy, F. R. Sheykhahmad, and N. Ghadimi, "A hybrid neural network–world cup optimization algorithm for melanoma detection," *Open Medicine*, vol. 13, no. 1, pp. 9–16, 2018.
- [9] A. A. A. El-Latif, M. S. Hossain, and N. Wang, "Score level multibiometrics fusion approach for healthcare," *Cluster Computing*, vol. 22, no. S1, pp. 2425–2436, 2019.
- [10] E. M. Abou-Nassar, A. M. Ilyyasu, P. M. El-Kafrawy, O.-Y. Song, A. K. Bashir, and A. A. A. El-Latif, "DITrust chain: towards blockchain-based trust models for sustainable healthcare IoT systems," *IEEE Access*, vol. 8, Article ID 111223, 2020.
- [11] A. A. Abd El-Latif, B. Abd-El-Atty, E. M. Abou-Nassar, and S. E. Venegas-Andraca, "Controlled alternate quantum walks based privacy preserving healthcare images in internet of things," *Optics and Laser Technology*, vol. 124, Article ID 105942, 2020.
- [12] C. Akyel and N. Arıcı, "LinkNet-B7: noise removal and lesion segmentation in images of skin cancer," *Mathematics*, vol. 10, no. 5, p. 736, 2022.
- [13] J. Yang, X. Sun, J. Liang, and P. L. Rosin, "Clinical skin lesion diagnosis using representations inspired by dermatologist criteria," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1258–1266, Salt Lake City, UT, USA, June 2018.
- [14] L. Liu, Y. Y. Tsui, and M. Mandal, "Skin lesion segmentation using deep learning with auxiliary task," *Journal of Imaging*, vol. 7, no. 4, p. 67, 2021.
- [15] Z. Mirikharaji, K. Abhishek, S. Izadi, and G. Hamarneh, "D-LEMA: Deep learning ensembles from multiple annotations-application to skin lesion segmentation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1837–1846, Nashville, USA, June 2021.
- [16] X. Wang, X. Jiang, H. Ding, Y. Zhao, and J. Liu, "Knowledge-aware deep framework for collaborative skin lesion segmentation and melanoma recognition," *Pattern Recognition*, vol. 120, Article ID 108075, 2021.
- [17] D. Dai, C. Dong, S. Xu et al., "Ms RED: a novel multi-scale residual encoding and decoding network for skin lesion segmentation," *Medical Image Analysis*, vol. 75, Article ID 102293, 2022.
- [18] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *Proceedings of the International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241, Springer, Munich, Germany, October 2015.
- [19] K. M. Hosny, M. A. Kassem, and M. M. Fouad, "Classification of skin lesions into seven classes using transfer learning with AlexNet," *Journal of Digital Imaging*, vol. 33, no. 5, pp. 1325–1334, 2020.

- [20] M. A. Kassem, K. M. Hosny, and M. M. Fouad, "Skin lesions classification into eight classes for ISIC 2019 using deep convolutional neural network and transfer learning," *IEEE Access*, vol. 8, Article ID 114822, 2020.
- [21] M. A. Kassem, K. M. Hosny, R. Damaševičius, and M. M. Eltoukhy, "Machine learning and deep learning methods for skin lesion classification and diagnosis: a systematic review," *Diagnostics*, vol. 11, no. 8, p. 1390, 2021.
- [22] T. Ridler and S. Calvard, "Picture thresholding using an iterative selection method," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 8, no. 8, pp. 630–632, 1978.
- [23] T. Wadhawan, N. Situ, K. Lancaster, X. Yuan, and G. Zouridakis, "SkinScan©: A portable library for melanoma detection on handheld devices," in *Proceedings of the 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 133–136, IEEE, Chicago, IL, USA, April 2011.
- [24] M. Silveira, J. C. Nascimento, J. S. Marques et al., "Comparison of segmentation methods for melanoma diagnosis in dermoscopy images," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 1, pp. 35–45, 2009.
- [25] A. Wong, J. Scharcanski, and P. Fieguth, "Automatic skin lesion segmentation via iterative stochastic region merging," *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 6, pp. 929–936, 2011.
- [26] V. Rajinikanth, N. S. Madhavaraja, S. C. Satapathy, and S. L. Fernandes, "Otsu's multi-thresholding and active contour snake model to segment dermoscopy images," *Journal of Medical Imaging and Health Informatics*, vol. 7, no. 8, pp. 1837–1840, 2017.
- [27] J. Fernandez Alcon, C. Ciuhu, W. Ten Kate et al., "Automatic imaging system with decision support for inspection of pigmented skin lesions and melanoma diagnosis," *IEEE journal of selected topics in signal processing*, vol. 3, no. 1, pp. 14–25, 2009.
- [28] M. Emre Celebi, H. A. Kingravi, H. Iyatomi et al., "Border detection in dermoscopy images using statistical region merging," *Skin Research and Technology*, vol. 14, no. 3, pp. 347–353, 2008.
- [29] Y. Yuan, M. Chao, and Y.-C. Lo, "Automatic skin lesion segmentation using deep fully convolutional networks with jaccard distance," *IEEE Transactions on Medical Imaging*, vol. 36, no. 9, pp. 1876–1886, 2017.
- [30] M. A. Al-Masni, M. A. Al-Antari, M.-T. Choi, S.-M. Han, and T.-S. Kim, "Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks," *Computer Methods and Programs in Biomedicine*, vol. 162, pp. 221–231, 2018.
- [31] L. Yu, H. Chen, Q. Dou, J. Qin, and P.-A. Heng, "Automated melanoma recognition in dermoscopy images via very deep residual networks," *IEEE Transactions on Medical Imaging*, vol. 36, no. 4, pp. 994–1004, 2017.
- [32] N. Nida, A. Irtaza, A. Javed, M. H. Yousaf, and M. T. Mahmood, "Melanoma lesion detection and segmentation using deep region based convolutional neural network and fuzzy C-means clustering," *International Journal of Medical Informatics*, vol. 124, pp. 37–48, 2019.
- [33] L. Bi, J. Kim, E. Ahn, A. Kumar, M. Fulham, and D. Feng, "Dermoscopic image segmentation via multistage fully convolutional networks," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 9, pp. 2065–2074, 2017.
- [34] Y. Tang, F. Yang, and S. Yuan, "A multi-stage framework with context information fusion structure for skin lesion segmentation," in *Proceedings of the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pp. 1407–1410, IEEE, Venice, Italy, April 2019.
- [35] F. Xie, J. Yang, J. Liu, Z. Jiang, Y. Zheng, and Y. Wang, "Skin lesion segmentation using high-resolution convolutional neural network," *Computer Methods and Programs in Biomedicine*, vol. 186, Article ID 105241, 2020.
- [36] Z. Yu, F. Jiang, F. Zhou et al., "Convolutional descriptors aggregation via cross-net for skin lesion recognition," *Applied Soft Computing*, vol. 92, Article ID 106281, 2020.
- [37] M. Attia, M. Hossny, S. Nahavandi, and A. Yazdabadi, "Skin melanoma segmentation using recurrent and convolutional neural networks," in *Proceedings of the 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, pp. 292–296, IEEE, Melbourne, Australia, April 2017.
- [38] Z. Ge, S. Demyanov, B. Bozorgtabar et al., "Exploiting local and generic features for accurate skin lesions classification using clinical and dermoscopy imaging," in *Proceedings of the 2017 IEEE 14th international symposium on biomedical imaging (ISBI 2017)*, pp. 986–990, IEEE, Melbourne, Australia, April 2017.
- [39] Y. Guo, A. S. Ashour, L. Si, and D. P. Mandalaywala, "Multiple convolutional neural network for skin dermoscopic image classification," in *Proceedings of the 2018 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, pp. 365–369, IEEE, Dubai, UAE, November 2018.
- [40] F. Rashid, A. Irtaza, N. Nida, A. Javed, H. Malik, and K. M. Malik, "Segmenting melanoma lesion using single shot detector (SSD) and level set segmentation technique," in *Proceedings of the 2019 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS)*, pp. 1–5, 2019.
- [41] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of big data*, vol. 6, no. 1, pp. 60–48, 2019.
- [42] C. F. Xception, "Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1251–1258, Honolulu, HI, USA, July 2017.
- [43] H. Li, X. He, F. Zhou et al., "Dense deconvolutional network for skin lesion segmentation," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 2, pp. 527–537, 2018.
- [44] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Re-thinking atrous convolution for semantic image segmentation," 2017, <https://arxiv.org/abs/1706.05587>.
- [45] L. Wei, k. Ding, and H. Hu, "Automatic skin cancer detection in dermoscopy images based on ensemble lightweight deep learning network," *IEEE Access*, vol. 8, pp. 99633–99647, 2020.
- [46] A. Rosebrock, *Deep Learning for Computer Vision with Python Starter Bundle*, PYIMAGESEARCH, Philadelphia, PA, USA, 1st edition, 2017.
- [47] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: surpassing human-level performance on imagenet classification," in *Proceedings of the IEEE international conference on computer vision*, pp. 1026–1034, Santiago, Chile, December 2015.
- [48] M. Q. Khan, A. Hussain, S. U. Rehman et al., "Classification of melanoma and nevus in digital images for diagnosis of skin cancer," *IEEE Access*, vol. 7, Article ID 90132, 2019.
- [49] M. Bukhari, K. B. Bajwa, S. Gillani et al., "An efficient gait recognition method for known and unknown covariate conditions," *IEEE Access*, vol. 9, pp. 6465–6477, 2021.

- [50] R. Ashraf, S. Afzal, A. U. Rehman et al., "Region-of-interest based transfer learning assisted framework for skin cancer detection," *IEEE Access*, vol. 8, Article ID 147858, 2020.
- [51] M. Bukhari, S. Yasmin, S. Sammad, A. A. Abd El-Latif, and A. Ahmed, "A deep learning framework for leukemia cancer detection in microscopic blood samples using squeeze and excitation learning," *Mathematical Problems in Engineering*, vol. 2022, Article ID 2801227, 18 pages, 2022.
- [52] D. Gutman, N. C. Codella, E. Celebi et al., "Skin lesion analysis toward melanoma detection: a challenge at the international symposium on biomedical imaging (ISBI) 2016, hosted by the international skin imaging collaboration (ISIC)," 2016, <https://arxiv.org/abs/1605.01397>.
- [53] B. Bozorgtabar, M. Abedini, and R. Garnavi, "Sparse coding based skin lesion segmentation using dynamic rule-based refinement," in *Proceedings of the international workshop on machine learning in medical imaging*, pp. 254–261, Springer, Athens, Greece, October 2016.

## Research Article

# Evaluating Histological Subtypes Classification of Primary Lung Cancers on Unenhanced Computed Tomography Based on Random Forest Model

Jianfeng Huang <sup>1</sup>, Wei He <sup>2</sup>, Haijia Xu <sup>3</sup>, Shan Yang <sup>1</sup>, Jiajun Dai <sup>1</sup>,  
Weifeng Guo <sup>1</sup> and Mengsu Zeng <sup>1</sup>

<sup>1</sup>Department of Radiology, Zhongshan Hospital, Fudan University, Shanghai 200032, China

<sup>2</sup>Department of Vascular Surgery, Zhongshan Hospital, Fudan University, Shanghai 200032, China

<sup>3</sup>School of Basic Medical Sciences, Fudan University, Shanghai 200032, China

Correspondence should be addressed to Weifeng Guo; [guo.weifeng@zs-hospital.sh.cn](mailto:guo.weifeng@zs-hospital.sh.cn) and Mengsu Zeng; [zeng\\_mengsu@163.com](mailto:zeng_mengsu@163.com)

Received 12 May 2022; Revised 7 July 2022; Accepted 21 January 2023; Published 6 February 2023

Academic Editor: Liwei Hu

Copyright © 2023 Jianfeng Huang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Lung cancer is the leading cause of cancer-related death in many countries, and an accurate histopathological diagnosis is of great importance in subsequent treatment. The aim of this study was to establish the random forest (RF) model based on radiomic features to automatically classify and predict lung adenocarcinoma (ADC), lung squamous cell carcinoma (SCC), and small cell lung cancer (SCLC) on unenhanced computed tomography (CT) images. Eight hundred and fifty-two patients (mean age: 61.4, range: 29–87, male/female: 536/316) with preoperative unenhanced CT and postoperative histopathologically confirmed primary lung cancers, including 525 patients with ADC, 161 patients with SCC, and 166 patients with SCLC, were included in this retrospective study. Radiomic features were extracted, selected, and then used to establish the RF classification model to analyse and classify primary lung cancers into three subtypes, including ADC, SCC, and SCLC according to histopathological results. The training (446 ADC, 137 SCC, and 141 SCLC) and testing cohorts (79 ADC, 24 SCC, and 25 SCLC) accounted for 85% and 15% of the whole datasets, respectively. The prediction performance of the RF classification model was evaluated by F1 scores and the receiver operating characteristic (ROC) curve. On the testing cohort, the areas under the ROC curve (AUC) of the RF model in classifying ADC, SCC, and SCLC were 0.74, 0.77, and 0.88, respectively. The F1 scores achieved 0.80, 0.40, and 0.73 in ADC, SCC, and SCLC, respectively, and the weighted average F1 score was 0.71. In addition, for the RF classification model, the precisions were 0.72, 0.64, and 0.70; the recalls were 0.86, 0.29, and 0.76; and the specificities were 0.55, 0.96, and 0.92 in ADC, SCC, and SCLC. The primary lung cancers were feasibly and effectively classified into ADC, SCC, and SCLC based on the combination of RF classification model and radiomic features, which has the potential for noninvasive predicting histological subtypes of primary lung cancers.

## 1. Introduction

Random forest (RF) algorithm, proposed by Leo Breiman [1], is an ensemble learning algorithm based on classification and regression trees (CART). The RF algorithm contains several CARTs, and each one is independent. Therefore, the RF algorithm performs insensitively to the overfitting problem of the training cohort and has superior noise immunity, which is not sensitive to default values [2]. The RF

algorithm is widely applied in various fields, such as life sciences because RF classification models are versatile, have high prediction accuracy, and provide additional information such as variable importance [3]. According to literature reports, RF algorithms have excellent performance in evaluating the progression, prognosis, and gene mutation expression of various diseases [4–10].

Lung cancer is the leading cause of cancer-related death in many countries, and the accurate histopathological

diagnosis is of great importance in subsequent treatment [11, 12]. In previous studies, the RF model was mostly applied to detecting lung cancer, the classification of benign and malignant pulmonary nodules, and the analysis of lung cancer prognosis [13–16]. However, for therapeutic purposes, primary lung cancers fall into three major subtypes: lung adenocarcinoma (ADC), lung squamous cell carcinoma (SCC), and small cell lung cancer (SCLC), and distinguishing among subtypes is still particularly challenging. In this study, our aim is to establish a classification model combining the RF algorithm and radiomic features of unenhanced CT images to classify the primary lung cancers into ADC, SCC, and SCLC and to evaluate the prediction performance.

## 2. Materials and Methods

**2.1. Study Population.** Nine hundred and twenty patients with histopathologically confirmed primary lung cancer from January 2013 to August 2018 at Zhongshan Hospital, Fudan University, were retrospectively studied. The inclusion criteria were: (1) diagnosis of ADC, SCC, or SCLC confirmed by puncture or surgical specimen; (2) preoperative CT examination within 2 weeks before surgery. The exclusion criteria were: (1) patients receiving other treatments such as chemotherapy and radiotherapy before surgery; (2) patients with lesion boundaries that were difficult to identify on CT images; (3) patients with inadequate quality images on CT; (4) patients with two or more histopathological subtypes of primary lung cancer; and (5) patients with the lesion less than 1 cm in diameter, avoiding partial volume effects. Eight hundred and fifty-two patients with primary lung cancer (525 ADC, 161 SCC, and 148 SCLC) were ultimately included in this study. All study procedures were approved by the Ethics Committee of Zhongshan Hospital, Fudan University.

**2.2. Protocol of Unenhanced Computed Tomography and Segmentation.** All patients had preoperative CT examinations performed within 2 weeks before the puncture or surgery under breath-hold conditions at the end of inspiration, from the thoracic inlet to the diaphragm, by experienced radiologists. The parameters of  $\mu$ CT 760 (Shanghai United Imaging Healthcare) were: tube voltage = 120 kV, tube current = 130 mA, slice thickness = 1 mm, and the parameters of LightSpeed 16 (GE Healthcare) were: tube voltage = 120–140 kV, tube current = 140 mA, and slice thickness = 1 mm. All image data were stored in DICOM format.

ITK-Snap software (version 3.6.0) was used to segment each layer of the tumor lesions on the CT images in all cases to obtain a three-dimensional region of interest (ROI) [17], which was output in mha format for analysis (see Figure 1). The histopathological results (ADC, SCC, or SCLC) of each case were matched to the segmentation results. All procedures of ROI segmentation were performed by two experienced radiologists and finally confirmed by a senior radiologist.

## 3. Establishment of Random Forest Classification Model

**3.1. Extraction of Radiomic Features.** PyRadiomics package implemented in Python was used to extract radiomic features [18] (see Figure 2). The radiomic features of both the original and wavelet denoised (db2 was set as the wavelet basis) images were extracted, such as shape-based features, firstorder statistics, the gray level co-occurrence matrix (GLCM), the gray level run length matrix (GLRLM), the neighboring gray tone difference matrix (NGTDM), the gray level dependence matrix (GLDM), and the gray level size zone matrix (GLSZM) (see Figure 2).

The features were first normalized to the range 0–1, and then the support vector machine (SVM) was used to filter the features (see Figure 2). The variance inflation factor (VIF) was used to detect the collinearity of features which made the area under the receiver operating characteristic (ROC) curve (AUC) of SVM classification greater than 0.5, and features with VIF less than or equal to 5 were selected (see Figures 2 and 3). The formula of VIF is given as follows:

$$VIF = \frac{1}{1 - R^2}. \quad (1)$$

The spatially uniform relevant features (Relieff) algorithm was used to further filter the features (see Figure 2), and the final retained radiomic features were summarized in Table 1 (see Table 1). The importance score of each feature for predicting the histopathological subtypes of primary lung cancer are shown in Figure 4 (see Figure 4).

**3.2. Random Forest Algorithm.** Selected radiomic features were used to establish and RF classification model with the following parameters: “n\_estimators” = 100; “max\_depth” = 11; “min\_samples\_split” = 2; and “min\_samples\_leaf” = 4. In this study, 85% of the whole data (724 in all, 446 ADC, 137 SCC, and 141 SCLC, respectively) were randomly divided into the training cohort which was used for feature selection as well as model fitting, and 5-fold cross-validation was used to validate in the training cohort, while the remaining 15% (128 in all, 79 ADC, 24 SCC, and 25 SCLC, respectively) were divided into the testing cohort for validation (see Table 2 and Figure 5).

**3.3. Statistical Analysis.** Statistical analysis was performed using SPSS software (Version 22.0) and Python 3.8.0 (NumPy packages). Categorical variables were presented as quantities (percentages) and compared using the chi-square test or Fisher’s exact test, while continuous variables were presented as the mean  $\pm$  SD if normally distributed, and compared using the Kruskal–Wallis H test because of the heterogeneity of variance. AUC, sensitivity, specificity, and accuracy were used to evaluate the predictive performance of the classification model. In addition, the F1 score was also used to evaluate the efficiency of the classification models. A two-tailed  $p$  value  $< 0.05$  was considered statistically significant.

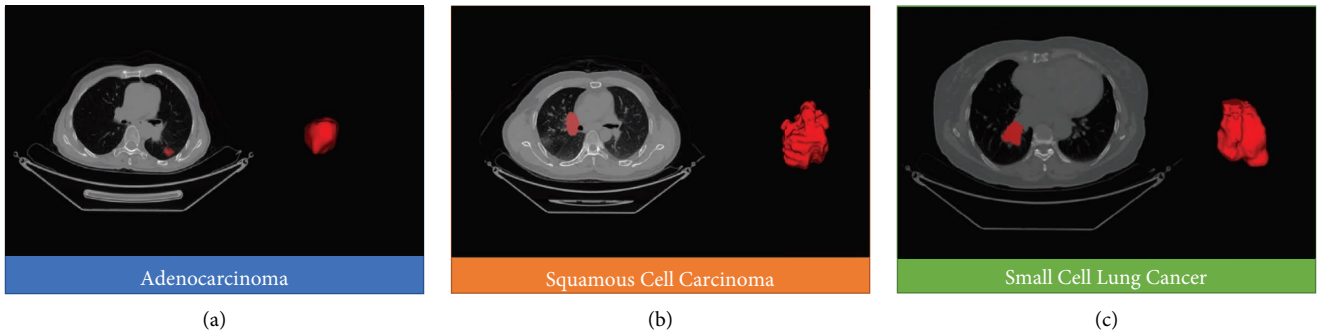


FIGURE 1: Segmentation of lesions on CT images and 3D ROI for (a) adenocarcinoma, (b) squamous cell carcinoma, and (c) small cell lung cancer.

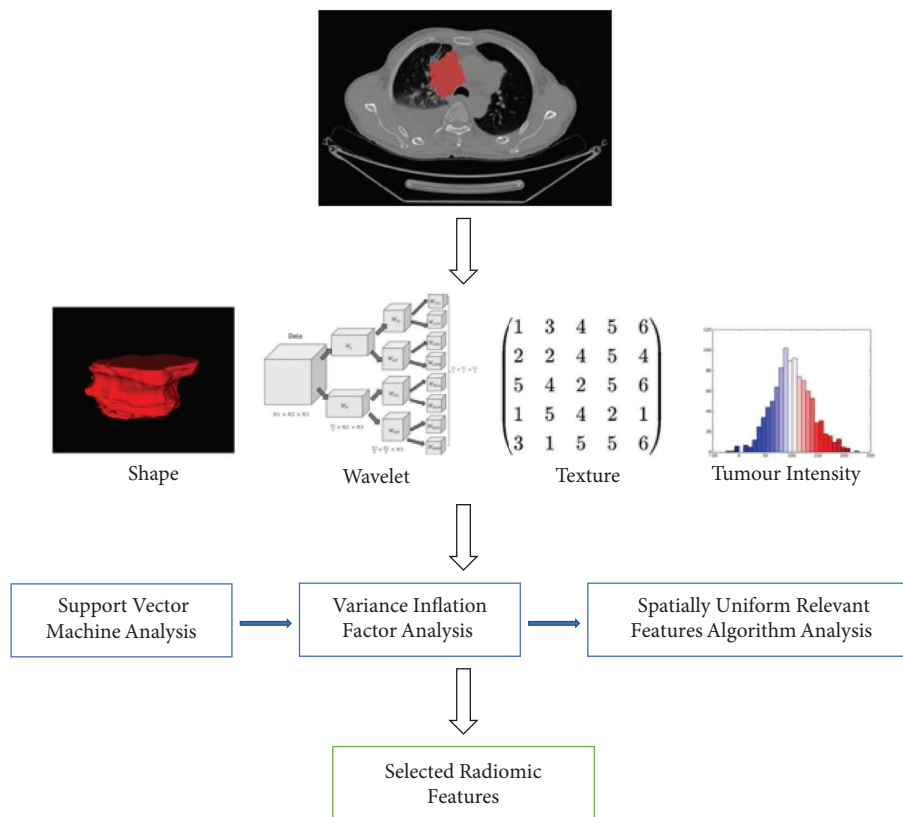


FIGURE 2: The flowchart for extraction and selection radiomic feature.

## 4. Results

**4.1. Patients' Clinical Characteristics.** Patients' clinical baselines and characteristics were summarized in Table 3 (see Table 3). Eight hundred and fifty-two patients with primary lung cancer (mean age: 61.4, range: 29–87, male/female: 536/316) were ultimately included in this study,

including 525 patients with ADC (61.6%, mean age: 60.4, range 29–87, male/female: 247/278), 161 patients with SCC (18.9%, mean age: 64.0, range 34–82, male/female: 148/13), and 166 patients with SCLC (19.5%, mean age: 62.1, range 38–82, male/female: 141/25), respectively, and randomly divided into the training cohort (724 patients) and testing cohort (128 patients) in the ratio of 85% to 15% (see Table 2).

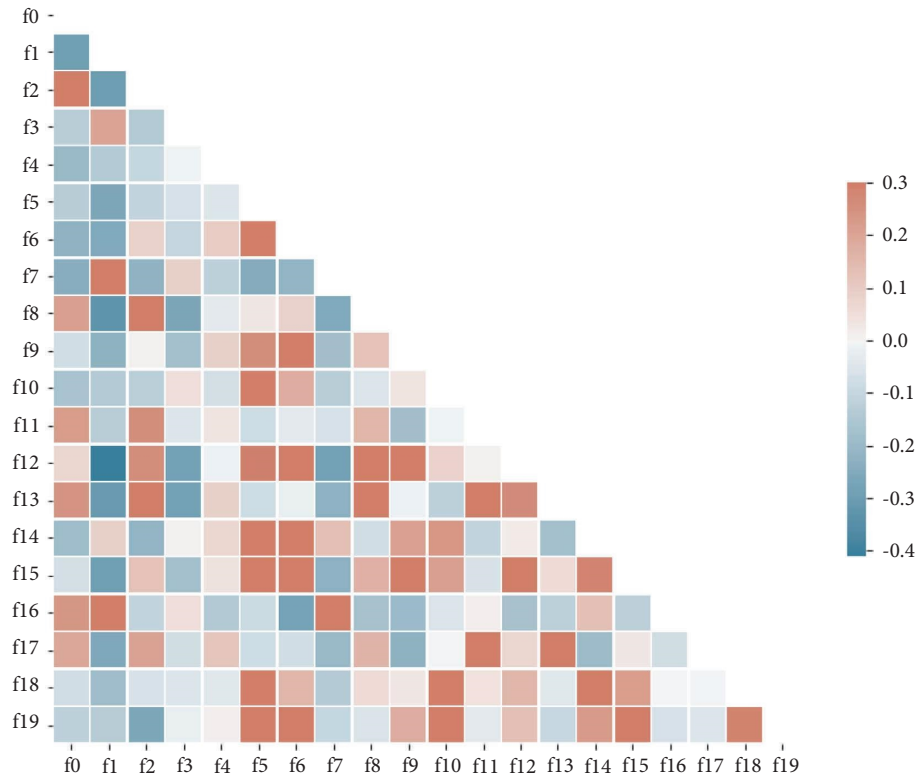


FIGURE 3: Correlation between radiomic features. 0 represents no correlation.

TABLE 1: Selected radiomic features.

No	Radiomics features
0	wavelet2-LLL_firstorder_RootMeanSquared
1	wavelet-HHH_ngtdm_Contrast
2	original_firstorder_InterquartileRange
3	wavelet2-HHH_glszm_SmallAreaLowGrayLevelEmphasis
4	wavelet2-HLH_firstorder_Mean
5	wavelet2-LHL_firstorder_Kurtosis
6	original_gldm_LargeDependenceHighGrayLevelEmphasis
7	wavelet2-LHH_glszm_SmallAreaLowGrayLevelEmphasis
8	wavelet-LHL_ngtdm_Complexity
9	original_glszm_GrayLevelNonUniformity
10	wavelet-LLH_firstorder_Kurtosis
11	original_ngtdm_Strength
12	wavelet-HLH_gldm_LargeDependenceHighGrayLevelEmphasis
13	wavelet2-HHH_glrlm_GrayLevelVariance
14	original_glrlm_RunVariance
15	wavelet2-LLH_gldm_LargeDependenceHighGrayLevelEmphasis
16	original_gldm_LowGrayLevelEmphasis
17	wavelet2-LLH_ngtdm_Strength
18	wavelet-LHH_firstorder_Kurtosis
19	original_firstorder_Kurtosis

Notably, the differences in age, gender, and TMN stage of patients among the three subtypes were statistically significant.

4.2. Predictive Performance of Random Forest Classification Model. Twenty radiomic features were ultimately selected after features were extracted and filtered from the

unenhanced CT to establish the RF classification model, including 7 firstorder features, 3 GLSZM features, 2 GLRLM features, 4 GLDM features, and 4 NGTDM features.

For the RF classification model, ROC-AUC was 0.74, 0.77, and 0.88 for the ADC, SCC, and SCLC, respectively, on the testing cohort. The average AUC for the three subtypes of classification was 0.80 (95% CI = 0.769–0.813) (see Figure 6).



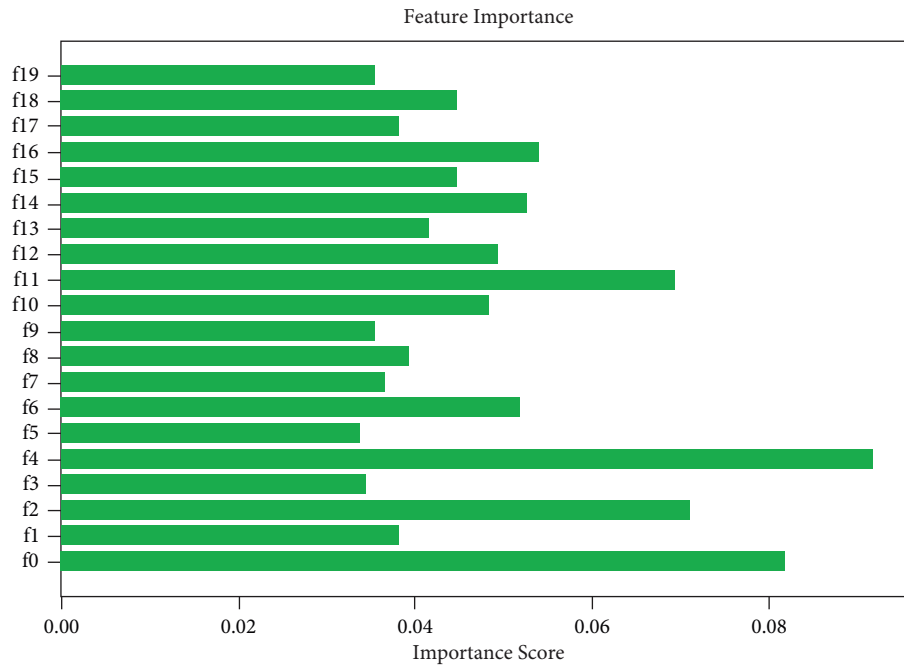


FIGURE 4: The importance score of each feature for predicting the histopathological subtypes of primary lung cancer.

TABLE 2: Cases in the training and testing cohort.

	Training ( $n = 724$ , 85%)	Testing ( $n = 128$ , 15%)
ADC	446	79
SCC	137	24
SCLC	141	25

In addition, the F1 scores achieved 0.80, 0.40, and 0.73 in ADC, SCC, and SCLC, respectively, and the weighted average F1 score was 0.71. Notably, the precisions were 0.72, 0.64, and 0.70, the recalls were 0.86, 0.29, and 0.76 and the specificities were 0.55, 0.96, and 0.92 in ADC, SCC, and SCLC (see Table 4).

## 5. Discussion

The histopathological diagnosis and classification of primary lung cancers are of great importance and crucial clinical value for the decision of optimal and individualized treatment schedules and the evaluation of prognosis [19]. In this study, RF algorithms combined with radiomic features on unenhanced CT images were used for noninvasive and preoperative prediction of subtype's classification of primary lung cancer. Radiomic features were extracted and filtered from enhanced CT images, and the ultimately 20 selected features were used to establish the RF classification model, which was trained and validated using the training cohort and the testing cohort. To be noted, 5-fold cross-validation was used for more accurate precision. Finally, the prediction performance of the model in classifying the three major subtypes (ADC, SCC, and SCLC) of primary lung cancers was evaluated.

The results showed that the RF classification model was able to accurately classify the three subtypes on the testing

cohort (AUC = 0.80). Particularly, the model performed better in predicting SCLC (AUC = 0.88) than ADC (AUC = 0.74) and SCC (AUC = 0.77). However, this model tended to misclassify SCC as ADC, thus the recalls of the RF model in ADC (0.86) and SCLC (0.76) were excellent, while inferior in SCC (0.29). It was probably because (1) the sample of SCC (161) was limited and much fewer than that of ADC (525) and (2) most of the SCC included in this study were central-type lung cancer, which was difficult to distinguish on the CT images, leading to inaccurate segmentation of ROI. Certainly, the reason for the misclassification deserved further investigation and verification. In this study, the selected 20 radiomic features were not the same as the features in previous studies (com\_radNet model), but they improved the predictive classification of SCLC [20].

Previously, a large number of studies have proven the excellent performance of the RF algorithm in classifying benign and malignant pulmonary nodules on CT and PET/CT [21–24]. Zhu et al. classified ADC and SCC in 129 patients with non-SCLC (NSCLC) based on 5 radiomic features with an AUC, specificity, and sensitivity of 0.89, 0.90, and 0.83, respectively, on the validation cohort [25]. Liu et al. classified 349 patients with NSCLC, including not only ADC and SCC but also large cell carcinoma and not otherwise specified based on radiomic features combined with SVM, and the classification accuracy was 0.86 on the testing set [26]. In this study, we expanded the samples and

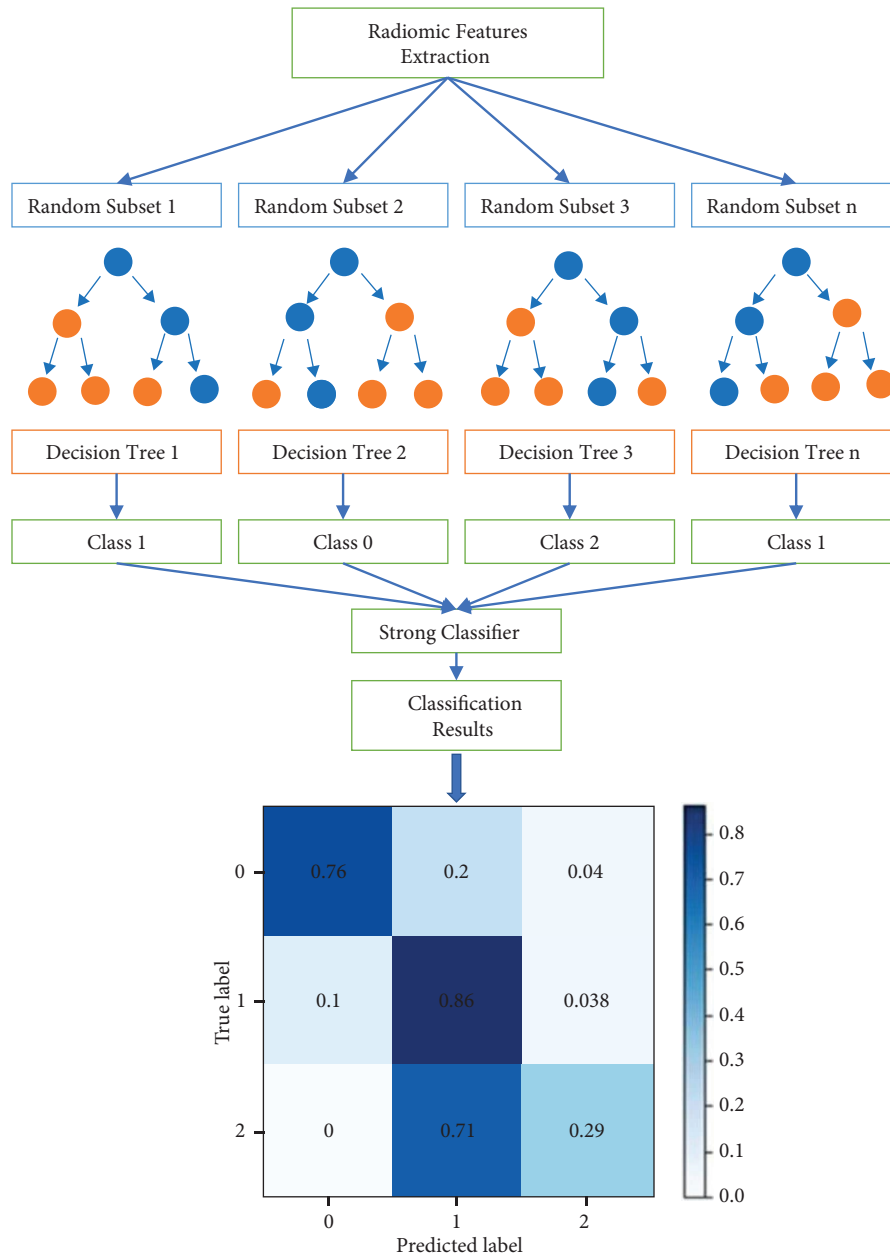


FIGURE 5: Flowchart of random forest algorithm.

TABLE 3: Demographics and characteristics.

	ADC ( $n = 525$ )	SCC ( $n = 161$ )	SCLC ( $n = 166$ )	$p$ values
Gender				<0.05
Male	247 (47.0%)	148 (91.9%)	141 (84.9%)	
Female	278 (53.0%)	13 (8.1%)	25 (15.1%)	
Age				<0.05
Mean $\pm$ SD	60.4 $\pm$ 10.5	64.0 $\pm$ 8.1	62.1 $\pm$ 9.5	
Range	18–87	23–82	41–86	
TNM				<0.05
I	139 (26.5%)	56 (34.8%)	12 (7.2%)	
II	70 (13.3%)	41 (25.5%)	19 (11.4%)	
III	91 (13.3%)	47 (29.2%)	57 (34.3%)	
IV	225 (42.9%)	17 (10.6%)	78 (47.0%)	

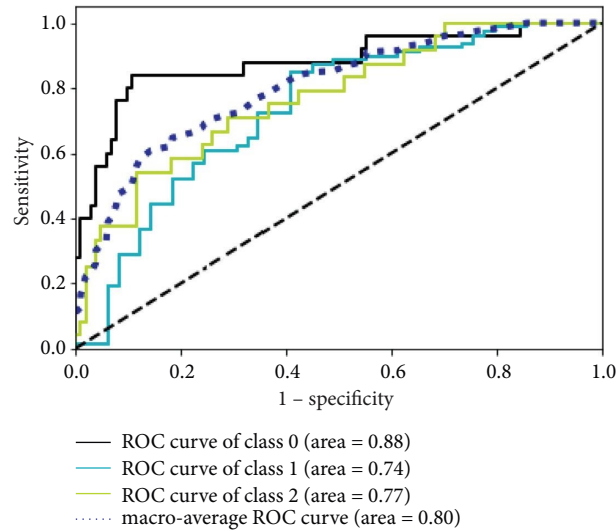


FIGURE 6: ROC curve analysis of the RF classification model. The black solid line represents SCLC (class 0), the blue solid line represents ADC (class 1), the green solid line represents SCC (class 2), and the dark blue dotted line represents the average.

TABLE 4: Prediction performance of the RF classification model.

Subtypes	F1 score	Precision	Recall	Specificity
ADC	0.80	0.76	0.86	0.55
SCC	0.40	0.64	0.29	0.96
SCLC	0.73	0.70	0.76	0.92

also investigated the classification between SCLC and NSCLC, with the considerably improved predictive performance of the RF classification model. To our knowledge, the only radiomics-based study on the classification of SCLC and NSCL identified ADC, SCC, and SCLC in a two-by-two comparison. The results showed good classification performance between ADC and SCLC (AUC=0.86) and between ADC and SCC (AUC=0.80) on unenhanced CT images and better performance on enhanced CT, but neither could effectively classify SCC and SCLC (AUC=0.62 and 0.66). To note, the RF classification model was able to classify SCLC with great performance in our study.

The study has some limitations. First, the sample of ADC was much larger than that of SCC and SCLC, mainly due to the different morbidities, which may affect the diagnostic performance of the model [27]. Furthermore, although we have excluded cases with blurred tumor borders, the possibility of missegmentation of nontumor tissues existed. Finally, large multicenter, prospective studies are essential for model expansion and optimization.

## 6. Conclusions

In conclusion, the noninvasive histopathological subtype classification of primary lung cancers has great clinical significance and value. In our study, the primary lung cancers were feasibly and effectively classified into ADC, SCC, and SCLC based on the combination of the RF classification model and radiomic features. Large studies are needed to optimize and validate the performance of the model. The RF classification model

combined with radiomic features on unenhanced CT images is able to provide additional information about patients and has the potential for clinical applications.

## Data Availability

The datasets used and/or analysed during the current study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest.

## Authors' Contributions

Jianfeng Huang and Wei He contributed equally to this paper.

## Acknowledgments

The authors would like to thank all participants in this study. This work was supported by the Shanghai Municipal Key Clinical Specialty (grant number shslczdzk03202).

## References

- [1] L. Breiman, "Random forest," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [2] H. Wang and G. Li, "A selective review on random survival forests for high dimensional data," *Quantitative Bio-Science*, vol. 36, no. 2, pp. 85–96, 2017.

- [3] W. G. Touw, J. R. Bayjanov, L. Overmars et al., "Data mining in the Life Sciences with Random Forest: a walk in the park or lost in the jungle?" *Briefings in Bioinformatics*, vol. 14, no. 3, pp. 315–326, 2013.
- [4] J. Li, Y. Tian, Y. Zhu et al., "A multicenter random forest model for effective prognosis prediction in collaborative clinical research network," *Artificial Intelligence in Medicine*, vol. 103, Article ID 101814, 2020.
- [5] D. Liu, X. Zhang, T. Zheng et al., "Optimisation and evaluation of the random forest model in the efficacy prediction of chemoradiotherapy for advanced cervical cancer based on radiomics signature from high-resolution T2 weighted images," *Archives of Gynecology and Obstetrics*, vol. 303, no. 3, pp. 811–820, 2021.
- [6] H. Patel, D. M. Vock, G. E. Marai, C. D. Fuller, A. S. R. Mohamed, and G. Canahuate, "Oropharyngeal cancer patient stratification using random forest based-learning over high-dimensional radiomic features," *Scientific Reports*, vol. 11, no. 1, Article ID 14057, 2021.
- [7] A. P. Cox, M. Raluy-Callado, M. Wang, A. M. Bakheit, A. P. Moore, and J. Dinot, "Predictive analysis for identifying potentially undiagnosed post-stroke spasticity patients in United Kingdom," *Journal of Biomedical Informatics*, vol. 60, pp. 328–333, 2016.
- [8] C. Buhnemann, S. Li, H. Yu et al., "Quantification of the heterogeneity of prognostic cellular biomarkers in ewing sarcoma using automated image and random survival forest analysis," *PLoS One*, vol. 9, Article ID e107105, 2014.
- [9] F. Miao, Y. P. Cai, Y. X. Zhang, Y. Li, and Y. T. Zhang, "Risk prediction of One-Year mortality in patients with cardiac arrhythmias using random survival forest," *Computational and Mathematical Methods in Medicine*, vol. 2015, pp. 1–10, Article ID 303250, 2015.
- [10] R. Sapir-Pichhadze and B. Kaplan, "Seeing the forest for the trees: random forest models for predicting survival in kidney transplant recipients," *Transplantation*, vol. 104, no. 5, pp. 905–906, 2020.
- [11] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2020," *CA: A Cancer Journal for Clinicians*, vol. 70, no. 1, pp. 7–30, 2020.
- [12] M. Santarpia, M. G. Daffina, N. Karachaliou et al., "Targeted drugs in small-cell lung cancer," *Translational Lung Cancer Research*, vol. 5, no. 1, pp. 51–70, 2016.
- [13] A. Jochems, I. El-Naqa, M. Kessler et al., "A prediction model for early death in non-small cell lung cancer patients following curative-intent chemoradiotherapy," *Acta Oncologica*, vol. 57, no. 2, pp. 226–230, 2018.
- [14] H. Shen, L. Chen, K. Liu et al., "A subregion-based positron emission tomography/computed tomography (PET/CT) radiomics model for the classification of non-small cell lung cancer histopathological subtypes," *Quantitative Imaging in Medicine and Surgery*, vol. 11, no. 7, pp. 2918–2932, 2021.
- [15] K. Mehta, A. Jain, J. Mangalagiri, S. Menon, P. Nguyen, and D. R. Chapman, "Lung nodule classification using biomarkers, volumetric radiomics, and 3D CNNs," *Journal of Digital Imaging*, vol. 34, no. 3, pp. 647–666, 2021.
- [16] U. Bashir, B. Kawa, M. Siddique et al., "Non-invasive classification of non-small cell lung cancer: a comparison between random forest models utilising radiomic and semantic features," *British Journal of Radiology*, vol. 92, no. 1099, Article ID 20190159, 2019.
- [17] P. A. Yushkevich, J. Piven, H. C. Hazlett et al., "User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability," *NeuroImage*, vol. 31, no. 3, pp. 1116–1128, 2006.
- [18] J. J. van Griethuysen, A. Fedorov, C. Parmar et al., "Computational radiomics system to decode the radiographic phenotype," *Cancer Research*, vol. 77, no. 21, pp. e104–e107, 2017.
- [19] J. W. Clark and D. L. Longo, "Recent progress in systemic treatment for lung cancer," *Current Opinion in Pulmonary Medicine*, vol. 24, no. 4, pp. 355–366, 2018.
- [20] Y. Guo, Q. Song, M. Jiang et al., "Histological subtypes classification of lung cancers on CT images using 3D deep learning and radiomics," *Academic Radiology*, vol. 28, no. 9, pp. e258–e266, 2021.
- [21] L. Rundo, R. E. Ledda, C. di Noia et al., "A Low-Dose CT-Based radiomic model to improve characterization and screening recall intervals of indeterminate prevalent pulmonary nodules," *Diagnostics*, vol. 11, no. 9, p. 1610, 2021.
- [22] W. Xing, H. Sun, C. Yan et al., "A prediction model based on DNA methylation biomarkers and radiological characteristics for identifying malignant from benign pulmonary nodules," *BMC Cancer*, vol. 21, no. 1, p. 263, 2021.
- [23] Y. Han, Y. Ma, Z. Wu et al., "Histologic subtype classification of non-small cell lung cancer using PET/CT images," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 48, no. 2, pp. 350–360, 2021.
- [24] W. Wenhao, H. Huihui, G. Jing, L. Xiaobing, H. Gang, and N. Shengdong, "Malignant-benign classification of pulmonary nodules based on random forest aided by clustering analysis," *Physics in Medicine and Biology*, vol. 64, no. 3, Article ID 035017, 2019.
- [25] X. Zhu, D. Dong, Z. Chen et al., "Radiomic signature as a diagnostic factor for histologic subtype classification of non-small cell lung cancer," *European Radiology*, vol. 28, no. 7, pp. 2772–2778, 2018.
- [26] J. Liu, J. Cui, F. Liu, Y. Yuan, F. Guo, and G. Zhang, "Multi-subtype classification model for non-small cell lung cancer based on radiomics: SLS model," *Medical Physics*, vol. 46, no. 7, pp. 3091–3100, 2019.
- [27] E. Linning, L. Lu, L. Li, H. Yang, L. H. Schwartz, and B. Zhao, "Radiomics for classifying histological subtypes of lung cancer based on multiphasic Contrast-Enhanced computed tomography," *Journal of Computer Assisted Tomography*, vol. 43, no. 2, pp. 300–306, 2019.