SpringerBriefs in Applied Sciences and Technology
Forensic and Medical Bioinformatics

Mohd Hafiz Arzmi · Anwar P. P. Abdul Majeed · Rabiu Muazu Musa · Mohd Azraai Mohd Razman · Hong-Seng Gan · Ismail Mohd Khairuddin · Ahmad Fakhri Ab. Nasir

Deep Learning in Cancer Diagnostics

A Feature-based Transfer Learning Evaluation

Springer

SpringerBriefs in Applied Sciences and Technology

Forensic and Medical Bioinformatics

The books of this series are submitted to SCOPUS, Google Scholar and Springerlink

Forensic and Medical Bioinformatics (FMB) series is a platform to bring interdisciplinary Technology driven Discovery and Disruptive content useful for Industry professionals, Researchers and Academicians and Advance Research students. We solicit contributions related to novel Research and innovation, Theory, applications, their technical elements and practical approaches related to

- Bio inspired technologies and Engineering in Biotechnology and Medical Sciences, Translational engineering
- Machine Learning and Artificial Intelligence in Health, Bioinformatics and Computational Biology
- Biosensors, Circuit and Systems Bio engineering, Communication technologies, Therapeutic systems and technologies, Biorobotics, Biomedical Signal Processing, BioMEMS, Neuro engineering, Biomedical circuits and systems
- Forensic Science and related laws, Case Studies and Scientific aids to investigation, Cyber Security, Bio Image Processing and Security

To submit your proposal or show your interest for this series please contact: amit.kumar@dnares.in OR loyola.dsilva@springer.com

Mohd Hafiz Arzmi · Anwar P. P. Abdul Majeed · Rabiu Muazu Musa · Mohd Azraai Mohd Razman · Hong-Seng Gan · Ismail Mohd Khairuddin · Ahmad Fakhri Ab. Nasir

Deep Learning in Cancer Diagnostics

A Feature-based Transfer Learning Evaluation

Springer

Mohd Hafiz Arzmi
Fundamental Dental and Medical Sciences
International Islamic University Malaysia
Kuantan, Pahang, Malaysia

Rabiu Muazu Musa
Center for Fundamental and Continuing Education, Department of Credited Co-curriculum
Universiti Malaysia Terengganu
Kuala Nerus, Terengganu, Terengganu, Malaysia

Hong-Seng Gan
School of AI and Advanced Computing
Xi'an Jiaotong—Liverpool University
Suzhou, China

Ahmad Fakhri Ab. Nasir
Faculty of Computing
Universiti Malaysia Pahang
Pekan, Malaysia

Anwar P. P. Abdul Majeed
School of Robotics
Xi'an Jiaotong—Liverpool University
Suzhou, China

Mohd Azraai Mohd Razman
Faculty of Manufacturing and Mechatronics Engineering Technology
Universiti Malaysia Pahang
Pekan, Malaysia

Ismail Mohd Khairuddin
Faculty of Manufacturing and Mechatronics Engineering Technology
Universiti Malaysia Pahang
Pekan, Pahang, Malaysia

ISSN 2191-530X ISSN 2191-5318 (electronic)
SpringerBriefs in Applied Sciences and Technology
ISSN 2196-8845 ISSN 2196-8853 (electronic)
SpringerBriefs in Forensic and Medical Bioinformatics
ISBN 978-981-19-8936-0 ISBN 978-981-19-8937-7 (eBook)
https://doi.org/10.1007/978-981-19-8937-7

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

I dedicate this book to my beloved parents, Arzmi Mansor and Safiah Abdul Aziz; my wife, Nurul Izzah Zulkifli; my children, Iffah Humaira Mohd Hafiz, Iffah Huriyya Mohd Hafiz, Iffah Huwayna Mohd Hafiz and Izzat Huzaifah Mohd Hafiz; my research group, Cluster of Cancer Research Initiative IIUM (COCRII), and also to my mentor, Anwar P. P. Abdul Majeed, for being very supportive until the very end.

Mohd Hafiz Arzmi

This book is dedicated to my parents, Ab. Nasir and Sukrinah, parents-in-law, Late Mohamad Taib and Late Azizah, my beloved wife, Farahiyah, as well as my cutie-pies, Ayshah, Hazeem, Qayyum and Ali.

Ahmad Fakhri Ab Nasir

The book is dedicated to my wife and son, Nurjalilah Yatim and Asif Azraai, to my mom Aida Ismail as well as my siblings, Azhar, Azmil, Azam, Azri and Azra. We are all gonna make it.

Mohd Azraai Mohd Razman

I would like to dedicate this book to my family, Mohd Khairuddin Shafie, Hayati Abu Bakar as well as my precious wife, Nursyahirah Zainal Abidin.

Ismail Mohd Khairuddin

I would like to dedicate this book to my parents, C. K. Gan, K. H. Tan as well as my sister, C. Y. Gan, and brother, H. K. Gan

Hong Seng Gan

This book is dedicated to my beloved family and friends. Let us keep thriving for excellence.

Rabiu Muazu Musa

I would like to dedicate this book to my wife Sharifah Maszura Syed Mohsin, my daughter, Saffiya Anwar and my son, Yusuff Anwar.

Anwar P. P. Abdul Majeed

Acknowledgements

We would like to acknowledge Prof. Dr. Zahari Taha for the guidance as well as the valuable suggestions for making the realisation of this book possible.

Mohd Hafiz Arzmi
Anwar P. P. Abdul Majeed
Ahmad Fakhri Ab Nasir
Mohd Azraai Mohd Razmaan
Ismail Mohd Khairuddin
Hong Seng Gan
Rabiu Muazu Musa

Contents

Chapter 1
Epidemiology, Detection and Management of Cancer: An Overview

Abstract Worldwide, breast cancer is recorded as the most prevalent cancer, with high mortality and morbidity rate of incidence amongst females. Meanwhile, lung cancer incidence, which is dominantly in the male population, was ranked as the third most common cancer in women. In addition, oral cancer, which refers to the cancers of the lip and oral cavity that include the lips, tongue, mouth, floor and palate, collectively had 90% of oral cancers begin as oral squamous cell carcinomas. In addition, skin cancer caused by the mutation of the skin cell has been ranked the seventeenth most common cancer worldwide, with new cases of non-melanoma skin cancer reported at 324,637. Even though many interventions have been done, the cases remain high worldwide due to late detection and poor management of the cancers. This chapter describes the epidemiology of cancers and the current method of detection and management of the diseases.

Keywords Epidemiology · Cancer Detection · Cancer Management

1.1 Introduction

According to Globocan 2020, the total number of new cancer cases was reported at 19,292,789 cases [1]. The cancer incidence rate was highest in Asia, followed by Europe and Northern America. Meanwhile, the cancer mortality rate was highest in Asia, Europe, Latin America and the Caribbean. The highest cancer was reported from breast cancer, followed by prostate and lung cancer. The incidence of breast cancer was reported at 47.8 cases for every 100,000, indicating the frequency of cancer remains high.

On the other hand, the incidence of lung cancer was reported at 22.4 cases per 100,000, showing that cancer cases remain high even though the management of lung cancer has been improvised. Even though the incidence of oral cancer is still considered low compared to other types of cancers, however, worldwide, 377,713 new cases were reported. Furthermore, skin cancer which was ranked the seventeenth highest number of new cases was reported to have total number mortality of 57,045 in 2020. Due to this high incidence of cancers, this chapter describes the epidemiology

M. H. Arzmi et al., *Deep Learning in Cancer Diagnostics*,
SpringerBriefs in Forensic and Medical Bioinformatics,
https://doi.org/10.1007/978-981-19-8937-7_1

of breast, lung, oral and skin cancers and their current detection and management methods.

1.2 Epidemiology, Detection and Management of Breast Cancer

Worldwide, breast cancer is recorded as the most prevalent cancer, with high mortality and morbidity rate of incidence amongst females [2]. In 2020, more than 2.3 million females globally were diagnosed with breast cancer, and about 80% of patients were individuals aged >50 [3]. Breast cancer contributes to 1 in 4 cancer cases and 1 in 6 cancer deaths, ranked first in incidences in 159 and mortality in 110 countries. Besides, breast cancer is also the fifth leading cause of cancer mortality worldwide, with 685,000 deaths [4]. Breast cancer accounted for 11.7% of total cancer accumulated cases in 2020 [4] and was a leading cause of most cancer-related deaths in women; as of 2021 alone, 281,550 (30% of the total 2021 cancer cases) women were diagnosed with breast cancer, and 43,600 (15% of the total death expected) were projected to die due to cancer, based on research conducted by the American Cancer Society in the United States [4]. It is expected to have an upward trend due to the Westernisation of lifestyles such as delayed pregnancies, reduced breastfeeding, low age of menarche, lack of physical activity and poor diet, and high and better cancer registration and detection [5].

During the 1980s to 1990s, breast cancer incidence rates increased dramatically, especially in Northern America, Oceania and also Europe [2], reflecting changes in the prevalence of risk factors along with increased detection through mammographic screening before the number of incidences dropped due to elimination usage of menopausal hormone therapy and due to numerous screening participation [6]. Starting in 2007, a few countries in Europe and Oceania, such as Denmark, the United States, Ireland and Scotland, recorded an increase in incidences of breast cancer cases along with countries in South America, Asia and African regions including high-income Asian countries, such as Japan and Republic of South Korea [7]. Changes in lifestyle, social, cultural and environment driven by capitalism and growth in the proportion of women in the industrial sector contributed to the prevalence of breast cancer risk factors. Mortality and morbidity in breast cancer also increased in sub-Saharan Africa from the mid-1990s to the mid-2010s in Malawi, Nigeria and Seychelles. They rapidly soared in Eastern Cape and Zimbabwe, mirroring their weak health infrastructure and poor survival outcome as most breast cancer cases reported at a late-stage presentation resulted in the improbability of prognosis [8].

In detecting breast cancer, mammography screening is the gold standard for breast cancer cases, especially in women in their 40 s. However, due to the disadvantages of the diagnostic technique for younger women, other techniques were developed to cater for women in their 20 s–30 s [3]. The recent evolvement of various techniques led to the easier measurement of biological procedures like an evaluation of

gene expression, molecular biology and cellular biochemistry using digital methods, tomosynthesis, CAD, nanoparticles and other methods like the radionuclide method that led to the detection of breast tumours devoid of using the in vivo method and directly using radiation contact [9]. Aside from futuristic approaches to be considered in breast cancer detection with hopes of producing more timely and accurate diagnoses such as breath biopsy, mammary ductoscopy, X-ray diffraction of hair, HER2 testing, magnetic resonance elastography and optical imaging [10]. The severity and the stages of cancer determine systemic chemotherapy, endocrine therapy or HER2-directed therapy. In managing the early stages of breast cancer, the treatment will focus on the patients' ER, PR and HER2 status, whilst for late-stage breast cancer, the receptor status and location of metastatic sites are the focus [11].

1.3 Epidemiology, Detection and Management of Lung Cancer

Lung cancer or bronchogenic carcinoma is an abnormal growth initiated from the lung parenchyma or within the bronchi [12]. According to the International Agency for Research on Cancer, lung cancer was estimated to be the third most subsequent cancer incidence worldwide in 2020. Within 5 years, the prevalence of lung cancer had reached about 2.6 million. Moreover, lung cancer mortality showed the highest of any other cancer, with an estimated 1.8 million cases recorded worldwide. Generally, lung cancer incidence has been predominantly in the male population and ranked as the third most common cancer in women. Globally, China was presented with the highest number of lung cancer cases, following up with the United States of America and India [1].

The classification of lung cancer can be distinguished based on the histological perspective, such as non-small cell lung carcinoma (NSCLC) and small-cell lung carcinoma (SCLC) [13]. The NSCLC is the most frequently diagnosed lung cancer, representing more than 50% of general cases [14]. Furthermore, lung cancer screening provides conclusive outcomes for early detection when surgical options are accessible to elevate the issue [15]. Recently, the National Lung Screening Trial (NLST) conducted a study on the advanced screening method, which led to a declining lung cancer mortality in individuals that have been screened with low-dose helical computed tomography (LDCT) in contrast with the common screening method, the chest radiography method in the high-risk population [16].

In treating and managing lung cancer, the procedures were according to the stages and classification of cancer. In NSCLC stages I, II and II, lobectomy or pneumonectomy with mediastinal lymph node sampling followed by adjuvant or non-adjuvant chemotherapy is the standard of preference. However, for stage IV, the current treatment is more focused on treating symptoms and improving survival, where the conditions are believed to be incurable. Besides surgery and chemotherapy, advanced methods have been proposed, such as targeted therapy and immunotherapy for

NSCLC cases. The treatments for SCLC diagnosis are similar to NSCLC; however, the differences are in the types of chemotherapy used [12].

1.4 Epidemiology, Detection and Management of Oral Cancer

Oral cancer collectively refers to the cancers of the lip and oral cavity, including the lips, tongue, mouth, floor and palate [17]. 90% of oral cancers begin as oral squamous cell carcinomas (OSCC) [18]. Other oral cavity cancers include salivary gland malignancies, sarcomas, malignant odontogenic tumours, melanoma and lymphoma. Survival rates vary in countries and depend on the stage of the diagnosis; in HDI countries, the prognosis is better compared to in low HDI countries [4].

The Globocan global cancer statistics estimate for incidence and mortality worldwide for 36 cancers in 185 countries reports that in 2020, lip and oral cavity cancers totalled 377 713 new cases, representing 2.0% of all cancer sites and recorded 177,757 new deaths [4, 17]. The statistics showed an increase in incidence and mortality compared to 2018 estimates [19], where 355,000 new diagnoses were made, and over 177,000 deaths were estimated. Furthermore, lip and oral cavity cancer had a 10.2 per 100,000 incidences and 5.7 per 100,000 mortality rate in males in lower HDI countries, especially in South Central Asian nations such as India, Sri Lanka and Pakistan, and in Melanesia (Papua New Guinea) [19]. The high incidence rate in these countries has been attributed to the betel nut chewing practice in these countries [19]. In countries with higher HDI, such as Western Europe and Australia/New Zealand, high incidence rates have been linked to smoking tobacco, drinking alcohol and HPV infection for oropharyngeal cancer and UV light exposure for lip cancers [19].

Oral cancer usually presents at a late stage, and early diagnosis for cancer prevention is a major issue in treatment [20]. Currently, oral screening for critical signs and symptoms can improve patient prognosis. Clinical examination of the oral cavity includes observation for lesions, ulcerated lesions, mobile teeth, numbness or bleeding [21]. Lesions that are differentiated from non-cancer are dysplastic and/or malignant can present as white or pigmented, proliferative or wart-like [21]. Extraction sockets that have not healed after 6 weeks also should be considered as possible carcinoma [22]. Clear lesions precede malignancy; hence, their early identification and removal can prevent the transformation of the lesion into a malignant stage [22]. Mortality and morbidity have been linked to socioeconomic factors, lack of public awareness and delays in primary health care [20]. There is a need, however, to aid the visual examination of lesions with screening aids to adequately differentiate between malignant and benign lesions [21].

To improve diagnosis, tissue-fluorescence imaging and optical coherence tomography, biosensors based on nanoparticles, DNA analysis and proteomics using saliva samples have been suggested as tools [21]. However, these tools remain yet to be

proven in clinical trials, and discrepancies between biopsy and diagnosis test results mean that these tools are not robust, practical or economical enough to be used in clinical diagnosis of oral cancer. A review of the screening programmes for early detection and prevention of oral cancer reports that visual examination helped to reduce the mortality rate in high-risk oral cancer patients. In contrast, using more advanced tools did not reduce mortality rates [20].

The first line of management of oral cancer lies with dentists, whose thorough oral cavity examination can detect premalignant lesions and early-stage oral cancer before treatment [23]. Once detected, lesions should be referred to an Oral and Maxillofacial Surgeon involved in Head and Neck Oncology [23]. The tissue is biopsied, analysed histopathologically and then scanned radiologically to stage the tumour. CT scans, MRI, ultrasound and positron emission tomography (PET) are common imaging tools used in this assessment. Once the evaluation of oral cancer type, stage and location has been made, the management plan can be drafted [23]. Primary treatment of oral cancer is surgery to remove the primary tumour, which can be followed by radiation therapy and chemotherapy, depending on the stage and histopathologic evaluation of the tumour [24]. This is all accompanied by post-operative care that aims to optimise nutrition, restore pulmonary, oral and respiratory function, and prevent further complications [23]. Successful application of these steps from surgery through aftercare is critical to the successful outcome of the patient. After treatment, the dentist resumes his role in routine surveillance of recurring or new lesions in the patient, as well as management of the patient's dental health post-treatment and other functional and aesthetic implications of oral cancer [23].

To conclude, oral cancer incidence and mortality rates have increased in the past two years, especially affecting countries with lower HDI due to poor diagnosis leading to early intervention. Prevention of oral cancer requires understanding the risk factors across different countries, which are associated with the lifestyle and norms in the country/region. Early detection of oral cancer happens at the dentist, where irregularities such as lesions in the oral cavity and persistent wounds are assessed and then diagnosed. Surgery is the normal treatment route, followed by radiation therapy and chemotherapy. Finally, post-operative care can play a major role in the successful return to the normal life of the patient.

1.5 Epidemiology, Detection and Management of Skin Cancer

According to WHO, a total of 324,635 new cases were reported from melanoma skin cancer, including nodular melanoma, superficial spreading melanoma, acral lentiginous and lentigo maligna [1]. The highest incidence was reported in Europe, followed by Northern America and Asia. Meanwhile, the mortality was reported to be the highest in Europe, followed by Asia. The number of new cases was higher in males than females, with 173,844 cases and 150,791 cases, respectively. Meanwhile,

the mortality rate was also higher in males than females, and the entire case reported was 32,385 and 24,658 death cases, respectively.

Most cancer cases lie under the umbrella of non-melanoma categories, such as squamous cell carcinoma, basal cell carcinoma and sebaceous gland carcinoma [25]. A total of 1,198,078 non-melanoma skin cancer cases were reported in 2020, with males exhibiting higher incidence than females [1]. Similarly, the mortality of males due to non-melanoma skin cancer was higher than that of females, with 37,596 and 26,135 death cases reported. Northern America had the highest case incidence, with 586,575 cases, followed by Europe and Asia. Meanwhile, the mortality case was reported to be highest in Asia, with 27,765 cases, followed by Europe and Africa.

Early diagnosis has been critical for skin cancer management [25]. Amongst the symptom of skin cancer includes the colour fadedness of the skin, ulceration and the increased size of the mole [26]. The biopsy method, which is a procedure of removing a sample from a suspected skin lesion for medical examination, has been the most common method for detecting skin cancer. This process is slow, painful and time-consuming.

1.6 Conclusion

In conclusion, since the mortality cases of cancer, particularly breast, lung and oral cancers, remain high worldwide; thus, there is an urgent need to introduce machine learning for the early detection of cancer.

References

1. World Health Organization (WHO) (2022) Global health estimates 2020: Deaths by cause, age, sex, by country and by region, 2000–2019. WHO; 2020. Accessed October 15, 2022. https://gco.iarc.fr/today/fact-sheets-cancers
2. Siegel RL, Miller KD, Fuchs HE, Jemal A (2021) Cancer statistics, 2021. CA: A Cancer J Clin 71(1): 7–33. https://doi.org/10.3322/caac.21654
3. Bhushan A, Gonsalves A, Menon JU (2021) Current state of breast cancer diagnosis, treatment, and theranostics. Pharmaceutics 13(5):1–24. https://doi.org/10.3390/pharmaceutics13050723
4. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F (2021) Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer J Clin 71(3): 209–249. https://doi.org/10.3322/caac.21660
5. Porter P (2008) Westernizing women's risk of breast cancer. N Engl J Med 358(1):213–216
6. Heer E, Harper A, Escandor N, Sung H, McCormack V, Fidler-Benaoudia MM (2020) Global burden and trends in premenopausal and postmenopausal breast cancer: A population-based study. Lancet Glob Health 8(8):e1027–e1037. https://doi.org/10.1016/S2214-109X(20)30215-1
7. Bray F, McCarron P, Parkin DM (2004) The changing global patterns of female breast cancer incidence and mortality. Breast Cancer Res 6(6):229–239. https://doi.org/10.1186/bcr932

8. Joko-Fru WY, Jedy-Agba E, Korir A, Ogunbiyi O, Dzamalala CP, Chokunonga E, Wabinga H, Manraj S, Finesse A, Somdyala N, Liu B, McGale P, Jemal A, Bray F, Parkin DM (2020) The evolving epidemic of breast cancer in sub-saharan Africa: Results from the African cancer registry network. Int J Cancer 147(8):2131–2141. https://doi.org/10.1002/ijc.33014
9. Mishra J, Kumar B, Targhotra M, Sahoo PK (2020) Advanced and futuristic approaches for breast cancer diagnosis. Future J Pharm Sci 6(1). https://doi.org/10.1186/s43094-020-00113-2
10. Mambou SJ, Maresova P, Krejcar O, Selamat A, Kuca K (2018) Breast cancer detection using infrared thermal imaging and a deep learning model. Sensors (Switzerland), 18(9). https://doi.org/10.3390/s18092799
11. Stanisławek A (2021) Breast cancer—epidemiology, risk factors, classification, prognostic markers, and current treatment strategies—an updated review. 1–30
12. Siddiqui F, Siddiqui AH (2021) Lung Cancer—StatPearls—NCBI Bookshelf. StatPearls publishing LLC. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK482357/?report=reader#_NBK482357_pubdet_
13. Schabath MB, Cote ML (2019) Cancer progress and priorities: Lung cancer. Cancer Epidemiol Biomark Prev 28(10):1563–1579. https://doi.org/10.1158/1055-9965.EPI-19-0221
14. Rami-Porta R, Bolejack V, Giroux DJ, Chansky K, Crowley J, Asamura H, Goldstraw P (2014) The IASLC lung cancer staging project: The new database to inform the eighth edition of the TNM classification of lung cancer. J Thorac Oncol 9(11): 1618–1624. https://doi.org/10.1097/JTO.0000000000000334
15. Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, Sick JD (2011) Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med 365(5):395–409. https://doi.org/10.1056/NEJMoa1102873
16. Schabath MB, Massion PP, Thompson ZJ, Eschrich SA, Balagurunathan Y, Goldof D, Gillies RJ (2016) Differences in patient outcomes of prevalence, interval, and screen-detected lung cancers in the CT arm of the national lung screening trial. PLoS ONE 11(8):e0159880. https://doi.org/10.1371/journal.pone.0159880
17. Miranda-Filho A, Bray F (2020) Global patterns and trends in cancers of the lip. Tongue and Mouth Oral Oncol 102:104551
18. El-Naggar AK, Chan JKC, Grandis JR, Takata T, Slootweg PJ (2017) WHO classification of tumours of the head and neck
19. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer J Clin 68(6): 394–424
20. Brocklehurst P, Kujan O, O'Malley LA, Ogden G, Shepherd S, Glenny AM (2013) Screening programmes for the early detection and prevention of oral cancer. Cochrane Database Syst Rev, (11)
21. Chakraborty D, Natarajan C, Mukherjee A (2019) Advances in oral cancer detection. Adv Clin Chem 91:181–200
22. Wong TSC, Wiesenfeld D (2018) Oral cancer. Aust Dent J 63:S91–S99
23. Gupta NK, Mahajan M, Hore A (2019) Management strategies for oral cancer subsites. In prevention, detection and management of oral cancer. IntechOpen
24. Carneiro-Neto JN (2017) Protocols for management of oral complications of chemotherapy and/or radiotherapy for oral cancer: Systematic review and meta-analysis current medicina oral patologia oral Y. Cir Bucal 22(1):e15
25. Dildar M, Akram S, Irfan M, Khan HU, Ramzan M, Mahmood AR, Mahnashi MH (2021) Skin cancer detection: A review using deep learning techniques. Int J Environ Res Public Health 18(10):5479
26. Qadir MI (2016) Skin cancer: Etiology and management. Pak J Pharm Sci 29(3)

Chapter 2
A VGG16 Feature-Based Transfer Learning Evaluation for the Diagnosis of Oral Squamous Cell Carcinoma (OSCC)

Abstract Oral Squamous Cell Carcinoma (OSCC) is the most prevalent type of oral cancer. Early detection of such cancer could increase a patient's survival rate by 83%. This chapter shall explore the use of a feature-based transfer learning model, i.e., VGG16 coupled with different types of conventional machine learning models, viz. Support Vector Machine (SVM), Random Forest as well as *k*-Nearest Neighbour (*k*NN) as a means to identify OSCC. A total of 990 evenly distributed normal and OSCC histopathological images are split into the 60:20:20 ratio for training, testing and validation, respectively. A testing accuracy of 93% was recorded via the VGG16-RF pipeline from the study. Consequently, the proposed architecture is suitable to be deployed as artificial intelligence-driven computer-aided diagnostics and, in turn, facilitate clinicians for the identification of OSCC.

Keywords Computer-Aided Diagnosis · Transfer Learning · Oral Cancer · OSCC

2.1 Introduction

It has been reported that oral cancer has the sixth highest occurrence amongst the different types of cancers [1]. The lack of early detection of this form of cancer contributes to a high mortality rate. Nonetheless, it is worth noting that Oral Squamous Cell Carcinoma (OSCC) accounts for more than 90% of oral cases. More often than not, smoking and the consumption of alcohol are deemed to be the main causes of OSCC [2]. As remarked earlier, early detection could allow patients to seek life-prolonging treatments. Traditional means of diagnosing such ailments by oncologists are rather labour-intensive, however, with the aid of computer-aided diagnostics (CAD) powered by artificial intelligence, the aforesaid predicament could be alleviated. Researchers have developed different deep learning architectures for diagnosing normal and malignant characteristics of oral cancer, especially OSCC.

Palaskar et al. [3] compared the efficacy of different transfer learning models, namely ResNet50, InceptionV3 and MobileNet against two conventional Convolutional Neural Network (CNN) models, dubbed as Large CNN and Small CNN in diagnosing histopathological OSCC images. Different sampling technique was

M. H. Arzmi et al., *Deep Learning in Cancer Diagnostics*,
SpringerBriefs in Forensic and Medical Bioinformatics,
https://doi.org/10.1007/978-981-19-8937-7_2

investigated to handle the imbalanced dataset. It was shown from the study that the InceptionV3 transfer learning model is able to classify better against the other evaluated models with a multi-site testing result of 83.66%.

In a recent study, Amin et al. [4] investigated the efficacy of concatenating different transfer learning models, i.e., VGG16, InceptionV3 and ResNet50 and comparing it with its individual models in classifying OSCC. A similar dataset employed by Palaskar et al. that was published by Rahman et al. [5] was used in the study. Hence, owing to the imbalanced dataset, the authors augmented it by oversampling. A total of 120 images was used as a testing dataset and it was shown from the study that the concatenated pipeline outperformed the individual models.

Abdul Rauf et al. [6] evaluated the performance of the InceptionV3 transfer learning model as a feature extractor whilst classifying OSCC images using Support Vector Machine (SVM), *k*-Nearest Neighbours (*k*NN) and Random Forest (RF) classifiers. It was shown from the study that the InceptionV3-RF pipeline demonstrates a test classification accuracy of 91%, in demarcating the normal and OSCC classes, suggesting that the proposed architecture has an attractive proposition. In this chapter, a feature-based transfer learning approach [7–9] by considering the VGG16 pre-trained CNN model with its fully connected layers replaced by the SVM, *k*NN and RF classifiers in the classification of OSCC is investigated.

2.2 Methodology

In this study, the histopathological images of the normal and OSCC tissue are obtained from an open-access repository curated by Rahman et al. [5]. It is worth noting that for the present study, the 400× magnified image set is used. This set contains 201 normal oral cavity images and 495 OSCC images. Owing to the imbalanced nature of the dataset, the normal images were oversampled to equate the total number of the OSCC images, therefore a total of 990 images was used in the study. A sample of the images from the two classes is shown in Fig. 2.1. The proposed architecture in the present investigation is a departure from a typical transfer learning pipeline, i.e., the fully connected layers are swapped with a conventional machine learning model [10, 11]. The features from the VGG16 pre-trained CNN model is then fed to the SVM, *k*NN and RF classifiers with its hyperparameters set to default from the scikit-learn library. The analysis was carried out using a Python IDE, namely Spyder 3.3.6, along with its associated Keras 2.3.1 and TensorFlow 1.14.0 libraries in evoking the VGG16 model. The performance of the pipelines is evaluated via the classification accuracies as well as the confusion matrix.

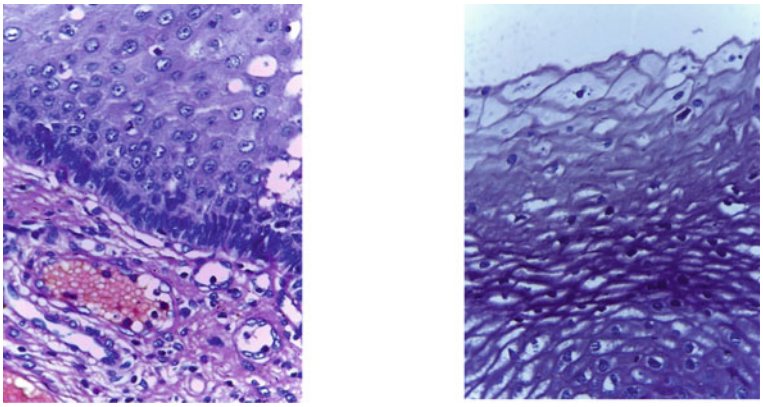


Fig. 2.1 Histopathological images for (**a**) Normal tissue (**b**) OSCC tissue

2.3 Results and Discussion

The performance of the evaluated pipelines is shown in Fig. 2.2. It is apparent that the VGG16-RF pipeline outperformed the other pipelines evaluated across all datasets. The VGG16-SVM pipeline performs reasonably well with a reduction of 3% in testing accuracy. The worst performing pipeline is the VGG16-*k*NN pipeline, suggesting that the default *k*NN classifier is unable to discern well the features extracted by the VGG16 model from the histopathological images. A similar observation with regards to the *k*NN model is reported in [6] from the features extracted via the InceptionV3 model. The confusion matrix on the test dataset of the evaluated pipelines is depicted in Fig. 2.3. The normal and OSCC classes are denoted as 0, and 1, respectively. It could be seen that no misclassification transpired on the N class across all pipelines. It could be seen that only five of the normal tissues are diagnosed as OSCC for the RF-based pipeline, whilst both the SVM and RF pipelines misdiagnosed six OSCC as normal. Considering the small fraction of misclassification, the proposed pipeline if deployed could significantly facilitate oncologists in the diagnosis of OSCC.

2.4 Conclusion

The chapter has demonstrated that the proposed architecture is able to discern the classes of normal and malignant oral tissue reasonably well, particularly the VGG16-RF pipeline. The deployment of such a model could facilitate oncologists in the diagnosis of OSCC, which is undoubtedly one of the most prevalent types of oral cancer.

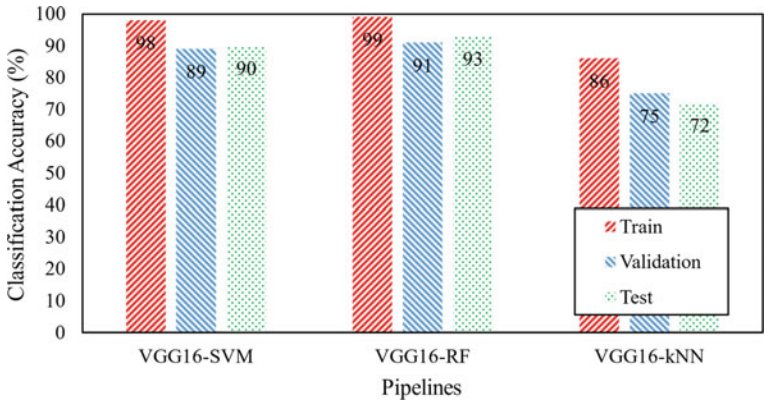


Fig. 2.2 The performance of the different VGG16 pipelines evaluated in terms of classification accuracy

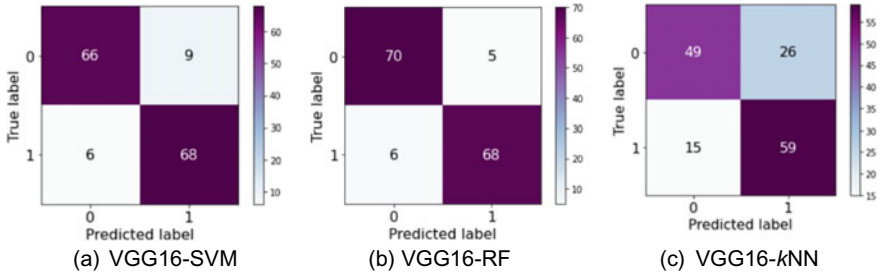


Fig. 2.3 The confusion matrix of pipelines on the testing dataset

References

1. Warnakulasuriya S (2009) Global epidemiology of oral and oropharyngeal cancer. Oral Oncol 45:309–316. https://doi.org/10.1016/j.oraloncology.2008.06.002
2. Hung LC, Kung PT, Lung CH, Tsai MH, Liu SA, Chiu LT, Huang KH, Tsai WC (2020) Assessment of the risk of oral cancer incidence in a high-risk population and establishment of a predictive model for oral cancer incidence using a population-based cohort in Taiwan. Int J Environ Res Public Health 17. https://doi.org/10.3390/ijerph17020665
3. Palaskar R, Vyas R, Khedekar V, Palaskar S, Sahu P (2020) Transfer learning for oral cancer detection using microscopic images. arXiv Prepr arXiv201111610
4. Amin I, Zamir H, Khan FF (2021) Histopathological image analysis for oral squamous cell carcinoma classification using concatenated deep learning models. medRxiv
5. Rahman TY, Mahanta LB, Das AK, Sarma JD (2020) Histopathological imaging database for oral cancer analysis. Data Br 29:105114 . https://doi.org/10.1016/j.dib.2020.105114
6. Abdul Rauf AR, Mohd Isa WH, Khairuddin IM, Mohd Razman MA, Arzmi MH, P. P. Abdul Majeed A (2022) The classification of oral squamous cell carcinoma (OSCC) by means of transfer learning. Lect Notes Networks Syst 429 LNNS:386–391. https://doi.org/10.1007/978-3-030-97672-9_34/FIGURES/5

7. Mohamed Ismail AH, Mohd Razman MA, Mohd Khairuddin I, Musa RM, Abdul Majeed APP (2021) The diagnosis of COVID-19 by means of transfer learning through X-ray Images. In: International Conference on Control, Automation and Systems. pp 592–595
8. Noor FNM, Isa WHM, Majeed APPA (2020) The diagnosis of diabetic retinopathy by means of transfer learning with conventional machine learning pipeline. MEKATRONIKA 2:62–67
9. Mat Jizat JA, Abdul Majeed APP, Ab. Nasir AF, Taha Z, Yuen E, Lim SX, (2022) Evaluation of the transfer learning models in wafer defects classification. Lect Notes Electr Eng 730:873–881
10. Abdullah MA, Ibrahim MAR, Shapiee MNA, Zakaria MA, Razman MAM, Musa RM, Osman NAA, Majeed APPA (2021) The classification of skateboarding tricks via transfer learning pipelines. PeerJ Comput Sci 7:e680
11. Noor FNM, Mohd Isa WH, Khairuddin IM, Razman MAM, Jizat JAM, Nasir AFA, Musa RM, P. P. Abdul Majeed A (2021) The diagnosis of diabetic retinopathy: A transfer learning with support vector machine approach. Adv Intell Syst Comput 1350 AISC:391–398

Chapter 3
The Classification of Breast Cancer: The Effect of Hyperparameter Optimisation Towards the Efficacy of Feature-Based Transfer Learning Pipeline

Abstract Breast cancer is the primary cause of death amongst women around the globe. The detection of such cancer at an early stage could allow patients to receive medical attention, which could increase their survival chances. This study attempts to investigate the effect of hyperparameter optimisation towards the classification efficacy of breast cancer. A total of 1080 histopathological images of benign and malignant breast tumours were split into the 70:15:15 ratio for training, testing and validation. The images are evaluated by two feature-based transfer learning pipelines, i.e., MobileNet-VanillaSVM and MobileNet-OptimisedSVM, respectively. It was demonstrated through this study that optimising the hyperparameters of the SVM model does increase the ability of the pipeline to discern the classes of the tumour.

Keywords Computer-Aided diagnosis · Transfer learning · Breast cancer · MobileNet · Hyperparameter optimisation

3.1 Introduction

The employment of automated computer-aided diagnostics (CAD) is non-trivial towards embracing precision medicine [1]. The World Cancer Report indicates that breast cancer is the leading cause of mortality and morbidity amongst women [2]. Nonetheless, early detection could provide patients with an avenue to seek appropriate medical treatment. Conventional means of detecting cancer via histopathological images by the pathologist are deemed to be laborious and prone to misdiagnosis [3]. Therefore, researchers have adopted different methods to address this predicament, and with the advent of artificial intelligence, the use of deep learning techniques has gained due attraction.

Han et al. [4] employed a structured deep learning model for the multi-classification of histopathological images. The BreaKHis and BreakHis with augmentation datasets were used in the study, in which eight classes of breast cancer were investigated. It was reported that the validation and testing dataset used real-world in-patient data. It was illustrated from the study that the proposed class structure-based deep convolutional neural network (CSDCNN) could achieve

M. H. Arzmi et al., *Deep Learning in Cancer Diagnostics*,
SpringerBriefs in Forensic and Medical Bioinformatics,
https://doi.org/10.1007/978-981-19-8937-7_3

appreciable accuracy for both image level and patient level against other models evaluated.

Different classical machine learning models against CNN were evaluated by Yadavendra and Chad [5] in classifying benign and malignant breast cancer tumours. The invasive ductal carcinoma (IDC) histopathology dataset was used in the study. The undersampling technique was used to mitigate the imbalanced dataset. Therefore, an equal distribution of 78,786 for benign and malignant tumours was subsequently split based on the 60:20:20 ratio for training, validation and testing, respectively. It was shown that the Xception-based CNN model could classify the classes rather well against other models evaluated.

Khan et al. [6] compared the efficacy of GooLeNet, VGGNet and ResNet against a proposed network that concatenated the three models for the classification of breast cytology images. The dataset that consists of 8000 images was obtained from the BreaKHis dataset as well as the locally developed dataset collected from the LRH hospital Peshwar, Pakistan. The 75:25 ratio was utilised for testing and training purposes. It was shown from the study that the proposed framework could achieve better accuracy in classifying the benign and malignant classes against the standalone pipelines.

Saber et al. [7] investigated the ability of different pre-trained convolution neural network (CNN) models, namely Inception V3, Inception-V2 ResNet, VGG16, VGG19, and ResNet50 in classifying three classes, viz. benign (B), malignant (M) and normal (N) from the mammographic image analysis-society (MIAS) dataset. The dataset consists of 61, 52 and 209 of B, M and N images, respectively. The 80:20 hold-out and tenfold cross-validation technique was used to develop and evaluate the transfer learning pipelines. It was shown from the study that the VGG16 pipeline is superior to other pipelines based on both cross-validation techniques investigated.

It could be observed from the aforesaid literature that the utilisation of deep learning techniques is rather attractive for the diagnosis of breast cancer. In this chapter, a feature-based transfer learning pipeline [8, 9], particularly the use of MobileNet, is appraised by considering the default or vanilla Support Vector Machine (SVM) classifier and Optimised SVM, i.e., the hyperparameters of the SVM model are optimised towards its ability in classifying benign and malignant breast cancers.

3.2 Methodology

In the present study, the BreaKHis 400X dataset curated by Spanhol et al. was used [10]. The microscopic biopsy images of benign (B) and malignant (M) breast tumours were undersampled to have an equal representation of 540 for each class. Figure 3.1 depicts the sample images for each class. The dataset was further split into the 70:15:15 ratio for training, testing and validation, correspondingly. It is worth noting at this juncture that the fully connected layer of the pre-trained CNN model is swapped with the Support Vector Machine (SVM) classifier. In this study, the MobileNet pre-trained CNN model is used to extract the information from the

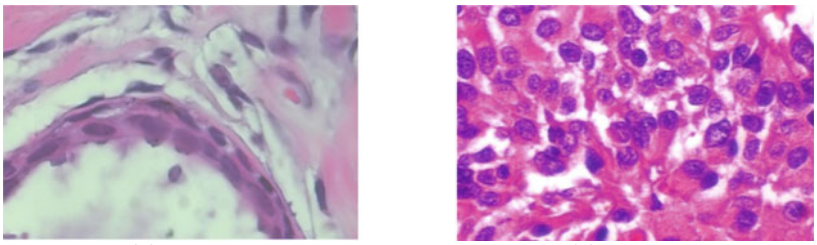


Fig. 3.1 Breast cancer histopathological image classes

images. It is then fed into a default SVM model, in which its hyperparameters were not optimised (default hyperparameters of the scikit-learn library) to classify the B and M classes. In order to investigate the effect of hyperparameter optimisation, the features were fed into the optimised SVM model, in which the kernel functions were varied between linear, cubic polynomial and radial basis functions, whilst the regularisation parameter, C, was varied between 0.01 and 100 with a tenfold interval. The grid search technique evoked with a fivefold cross-validation was used to determine the best hyperparameters [11]. Both the performance of the 'Vanilla' SVM and Optimised SVM were compared in terms of classification accuracy and the confusion matrix. The analysis was carried out using a Python IDE, namely Spyder 3.3.6, along with its accompanying Keras and TensorFlow libraries to evoke the deep learning models.

3.3 Results and Discussion

The performance of the different MoblieNet pipelines is shown in Fig. 3.2. It could be seen that by optimising the hyperparameters of the SVM model, the classification accuracy of both the training and testing dataset is improved. A linear kernel with a C value of 0.01 obtained through the grid search technique was deemed to provide better performance. Nevertheless, it is also worth noting that the lightweight MobileNet could also provide appropriate features from the histopathological images for the classifiers evaluated.

The confusion matrix of the evaluated pipelines on the test dataset is depicted in Fig. 3.3. The B and M classes are denoted as 0 and 1, respectively. It is apparent that the misclassification of the B as M is reduced by optimising the SVM's hyperparameters. It is worth noting that only two of the three hyperparameters are varied; therefore, a further improvement of the pipeline could be achieved by varying the gamma parameter. Nonetheless, it ought to be mentioned at this juncture that the proposed framework, upon judicious investigation, could further facilitate the adoption of CAD as a tool for precision medicine.

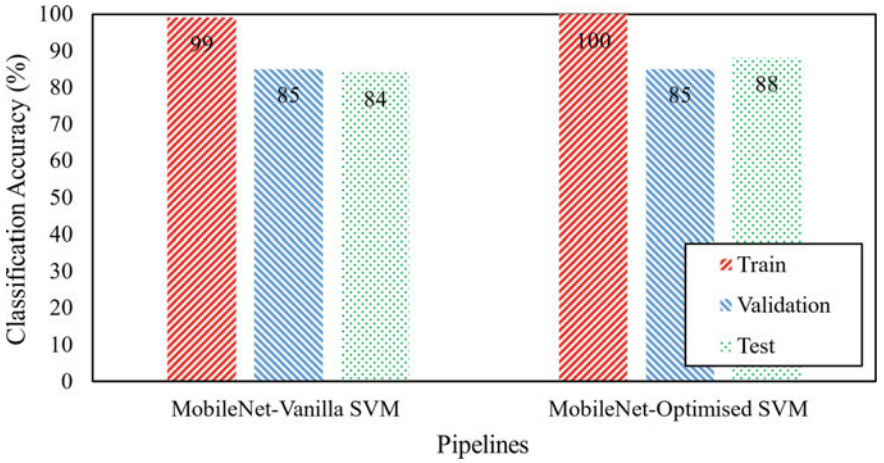
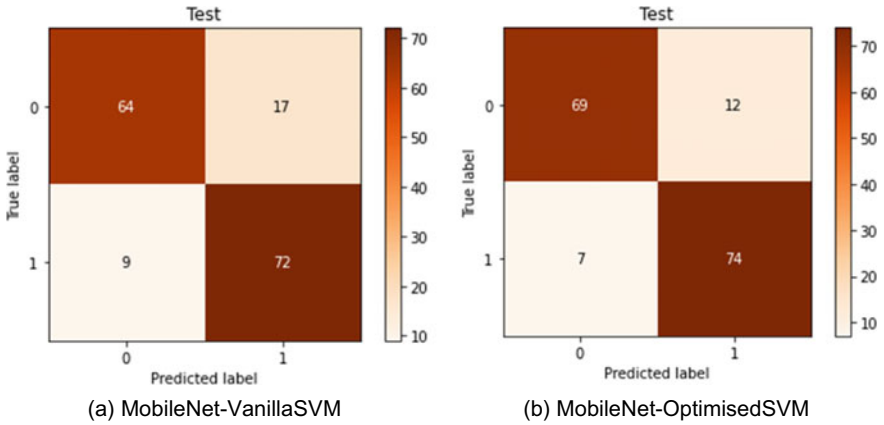


Fig. 3.2 The performance of the different MobileNet pipelines was evaluated in terms of classification accuracy



(a) MobileNet-VanillaSVM (b) MobileNet-OptimisedSVM

Fig. 3.3 The confusion matrix of the evaluated pipelines on the testing dataset

3.4 Conclusion

The study has illustrated that the accuracy of a given transfer learning pipeline could be further increased by optimising the hyperparameter of the swapped, fully connected classifier. The proposed framework, upon improvement, could facilitate pathologists in discerning the malignancy of breast cancer.

References

1. Collins FS, Varmus H (2015) A new initiative on precision medicine. N Engl J Med 372(9):793–795. https://doi.org/10.1056/nejmp1500523
2. Singh S, Kumar R (2022) Breast cancer detection from histopathology images with deep inception and residual blocks. Multimed Tools Appl 81:5849–5865. https://doi.org/10.1007/S11042-021-11775-2
3. Wang P, Hu X, Li Y, Liu Q, Zhu X (2016) Automatic cell nuclei segmentation and classification of breast cancer histopathology images. Signal Process 122:1–13. https://doi.org/10.1016/j.sigpro.2015.11.011
4. Han Z, Wei B, Zheng Y, Yin Y, Li K, Li S (2017) Breast cancer multi-classification from histopathological images with structured deep learning model. Scientific Reports 2017 7:1 7:1–10. https://doi.org/10.1038/s41598-017-04075-z
5. Chand S (2020) A comparative study of breast cancer tumor classification by classical machine learning methods and deep learning method. Mach Vis Appl 31:1–10
6. Khan SU, Islam N, Jan Z, Ud Din I, Rodrigues JJPC (2019) A novel deep learning based framework for the detection and classification of breast cancer using transfer learning. Pattern Recognit Lett 125:1–6. https://doi.org/10.1016/J.PATREC.2019.03.022
7. Saber A, Sakr M, Abo-Seida OM, Keshk A, Chen H (2021) A novel deep-learning model for automatic detection and classification of breast cancer using the transfer-learning technique. IEEE Access 9:71194–71209
8. Kumar JLM, Rashid M, Musa RM, Razman MAM, Sulaiman N, Jailani R, Abdul Majeed APP (2021) The classification of EEG-based wink signals: A CWT-Transfer learning pipeline. ICT Express. https://doi.org/10.1016/j.icte.2021.01.004
9. Mahendra Kumar JL, Rashid M, Muazu Musa R, Mohd Razman MA, Sulaiman N, Jailani R, P.P. Abdul Majeed A, (2021) The classification of EEG-based winking signals: a transfer learning and random forest pipeline. PeerJ 9:e11182. https://doi.org/10.7717/peerj.11182
10. Spanhol FA, Oliveira LS, Petitjean C, Heutte L (2016) A dataset for breast cancer histopathological image classification. IEEE Trans Biomed Eng (TBME) 63:1455–1462. https://doi.org/10.1109/tbme.2015.2496264
11. Abdullah MA, Ibrahim MAR, Shapiee MNA, Zakaria MA, Razman MAM, Musa RM, Osman NAA, Majeed APPA (2021) The classification of skateboarding tricks via transfer learning pipelines. Peer J Comput Sci 7:e680

Chapter 4
The Classification of Lung Cancer: A DenseNet Feature-Based Transfer Learning Evaluation

Abstract In the present study, a class of deep learning approaches is used to classify non-small-cell lung cancers. A total of 400 computed tomography (CT) images of lung cancer that are demarcated into normal, large cell carcinoma, adenocarcinoma and squamous cell carcinoma are split into the 70:15:15 ratio for training, testing and validation. The images are evaluated on different DenseNet-Support Vector Machine (SVM) pipelines, i.e., DenseNet121-SVM, DenseNet169-SVM and DenseNet201-SVM, respectively. It was shown from the present investigation that the DenseNet121-SVM pipeline is able to yield a test classification accuracy of 87%. Therefore, it could be demonstrated that the proposed architecture is able to classify the different variations of non-small-cell lung cancers reasonably well and could further facilitate the diagnosis of lung cancer by clinicians.

Keywords Computer-Aided diagnosis · Transfer learning · Lung cancer

4.1 Introduction

It has been reported that lung cancer contributes to the third most mortality amongst the different types of cancers around the globe [1]. Moreover, passive and active smokers constitute approximately 90% of cancer patients [2]. More often than not, patients are diagnosed with such cancer at a later stage, i.e., Stage Three or Four. Nonetheless, early detection could provide patients with an opportunity to seek medical treatment [3]. Conventional means of diagnosis would require clinicians to go through the computed tomography (CT) images scanned. However, such rudimentary methods are laborious and prone to misdiagnosis. Owing to the advancement of computational technology and artificial intelligence, particularly deep learning, the aforementioned issues could be circumvented via computer-aided diagnostics (CAD). Researchers, by large, have attempted to exploit different deep learning architectures to diagnose lung cancer.

The diagnosis of Solitary Pulmonary Nodules (SPN) by means of a variation of pre-trained convolution neural network (CNN) models was investigated by Apostolopoulos et al. [4]. Four different datasets (augmented and combined) that consist

M. H. Arzmi et al., *Deep Learning in Cancer Diagnostics*,
SpringerBriefs in Forensic and Medical Bioinformatics,
https://doi.org/10.1007/978-981-19-8937-7_4

of two classes, viz. benign and malignant nodules, were evaluated. The tenfold cross-validation technique was employed in the study. It was shown from the study that the VGG16 as a feature extractor, whilst the fully connected layers are connected to a Softmax activation function, yielded the best classification accuracy across all datasets against other evaluated models.

Ashhar et al. [2] investigated different pre-trained CNN models, namely GoogleNet, SqueezeNet, DenseNet, ShuffleNet and MobileNetV2, in order to classify CT-based lung tumour images into malignant and benign categories. The dataset was obtained from the Lung Image Database Consortium (LIDC-IDRI), which comprises 1423 malignant and 223 benign images. The dataset was split into 70% for training and 30% for testing. It was shown from the study that a training accuracy of 94.53% is achievable via GoogleNet to demarcate the two classes.

The classification of lung carcinoma by means of pre-trained CNN models, viz. VGG16, VGG19 and Xception were carried out by Humayun et al. [1]. A customised dataset was obtained from the Iraq-Oncology Teaching Hospital/National Center for Cancer Diseases (IQ-OTH/NCCD) that is demarcated into three categories, i.e., normal, benign and malignant. About 75% of the dataset was considered as the training set, whilst 25% as the testing set. A testing accuracy of 83.39%, 80.97% and 89.68% was reported for the VGG 16, VGG 19 and Xception pipelines, respectively, suggesting appreciable demarcation of the classes could be made.

Bębas et al. investigated different machine learning models, namely Support Vector Machine (SVM), k-Nearest Neighbours (kNN), Naïve Bayes (NB), Classification and Regression Trees (CART), Random Forest (RF) and Deep Learning architecture, to classify adenocarcinoma and squamous cell carcinoma [5]. The features were extracted from different image histogram analyses, for instance, histogram of oriented gradients (HOG) amongst others. It was shown that the HOG-SVM pipeline demonstrated the best performance with a classification accuracy of 75.48%.

It could be seen from the brief literature reported that, to a certain extent, the employment of machine learning and deep learning has somewhat provided a positive impact on computer-aided diagnostics. In the present investigation, a feature-based transfer learning approach [6, 7], particularly by exploiting different DenseNet models, i.e., DenseNet121, DenseNet169 and DenseNet201, is evaluated for its efficacy in the classification of non-small-cell lung cancers such as adenocarcinoma, squamous cell cancer and large cell carcinoma.

4.2 Methodology

The dataset was obtained from an open-access repository, viz. Kaggle [8]. A total of 400 computed tomography (CT) images of lung cancer are demarcated into normal (N), large cell carcinoma (LCC), adenocarcinoma (A) and squamous cell carcinoma (SCC). The CT images of the aforementioned classes are depicted in Fig. 1a to 1d. It is worth noting that undersampling was carried out in order to have an equal representation of the types of cancer. This is non-trivial, particularly in mitigating

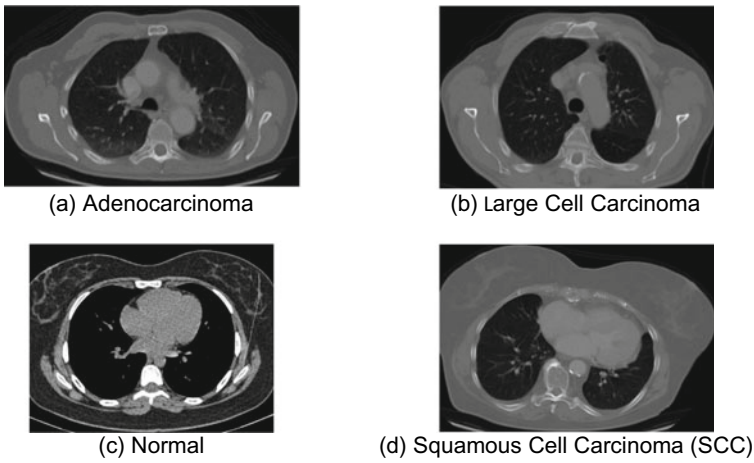


Fig. 4.1 Different lung cancer classification

the potential overfitting and biased representation of the data. The dataset was split into the 70:15:15 ratio for training, testing and validation, respectively. In the present study, the DenseNet family, i.e., DenseNet121, DenseNet169 and DenseNet201, were employed to extract the features from the CT-based images. It is worth noting that the proposed architecture in the present investigation is a departure from the conventional means of deploying a Convolutional Neural Network (CNN) model, i.e., the fully connected layers are swapped with a conventional machine learning model [9, 10]. The features were then fed into a Support Vector Machine (SVM) classifier with its hyperparameters used are default from the scikit-learn library. The analysis was carried out using a Python IDE, namely Spyder 3.3.6, along with its associated Keras and TensorFlow libraries. The performance of the pipelines is evaluated via the classification accuracies, as well as the macro average of the precision, recall and f1-scores, respectively.

4.3 Results and Discussion

Figure 4.2 depicts the performance of the evaluated pipelines. It could be seen that the DenseNet121-SVM pipeline yield a better accuracy with respect to the train, validation and test dataset against the other two evaluated pipelines. Although the DenseNet201-SVM pipeline demonstrated an equal test classification accuracy, nonetheless, it demonstrated a relatively poor validation accuracy, suggesting that the pipeline is not as robust as the DenseNet121-SVM pipeline. It is evident from the present study that the DenseNet121 model is able to extract the features of the CT images better in comparison to the other two DenseNet models.

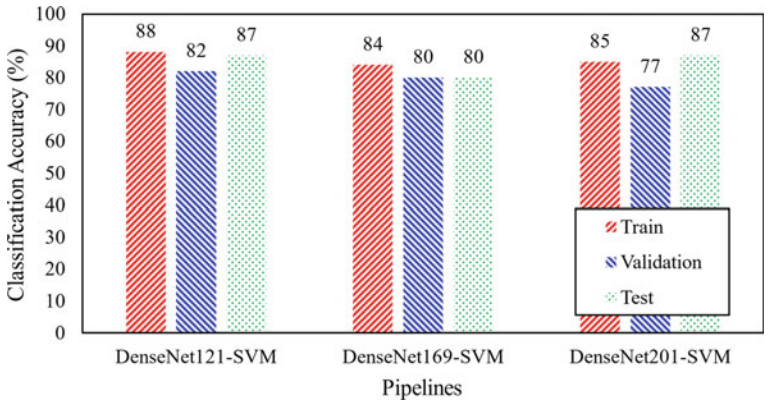
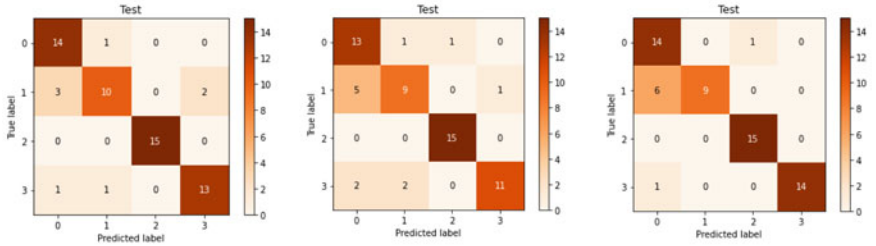


Fig. 4.2 The performance of the different DenseNet pipelines evaluated in terms of classification accuracy

Table 4.1 lists the macro average of the precision, recall and f1-scores, respectively, for the pipelines evaluated from the validation dataset. It is apparent from the performance measures that the DenseNet121-SVM pipeline is much superior compared to the other evaluated pipelines. The confusion matrix on the test dataset of the evaluated pipelines is shown in Fig. 4.3. The classes A, LCC, N and SCC are denoted as 0, 1, 2 and 3, respectively. It could be seen that no misclassification transpired on the N class across all pipelines, whilst only one misclassification transpired for the A class for the DenseNet121- and DenseNet201-SVM pipelines. A better classification of the LCC is observed by the DenseNet121-SVM pipeline. It is also apparent that the misclassification of this class is observed across all pipelines and is often confused with A. The occurrence of SCC being misclassified as A is also noticeable. Nevertheless, it is worth reiterating again that there was no misclassification that occurred between cancer and non-cancer classes, which is non-trivial to computer-aided diagnostics that would, in turn, alleviate the burden of radiologists and clinicians.

Table 4.1 Other performance measures on the validation dataset

Pipeline	Performance Measures (%)		
	Precision	Recall	F1-Score
DenseNet121-SVM	83	82	82
DenseNet169-SVM	84	80	80
DenseNet201-SVM	83	77	76



(a) DenseNet121-SVM (b) DenseNet169-SVM (c) DenseNet201-SVM

Fig. 4.3 The confusion matrix of pipelines on the testing dataset

4.4 Conclusion

The present study has demonstrated that the feature-based transfer learning approach, particularly, DenseNet121-SVM pipeline, is able to demarcate the different classes of lung cancer reasonably well. The proposed approach is deemed to be suitable for the diagnosis of lung cancer and would ultimately alleviate the burden of clinicians, especially radiologists, in their daily endeavours.

References

1. Humayun M, Sujatha R, Almuayqil SN, Jhanjhi NZ (2022) A transfer learning approach with a convolutional neural network for the classification of lung carcinoma. In: Healthcare. MDPI, p 1058
2. Ashhar SM, Mokri SS, Abd Rahni AA, Huddin AB, Zulkarnain N, Azmi NA, Mahaletchumy T (2021) Comparison of deep learning convolutional neural network (CNN) architectures for CT lung cancer classification. Int J Adv Technol Eng Explor 8:126
3. Kan Chan Siang MD, John CKM (2016) A review of lung cancer research in Malaysia. Med J Malaysia 71:71
4. Apostolopoulos ID, Pintelas EG, Livieris IE, Apostolopoulos DJ, Papathanasiou ND, Pintelas PE, Panayiotakis GS (2021) Automatic classification of solitary pulmonary nodules in PET/CT imaging employing transfer learning techniques. Med Biol Eng Comput 59:1299–1310
5. Bębas E, Borowska M, Derlatka M, Oczeretko E, Hładuński M, Szumowski P, Mojsak M (2021) Machine-learning-based classification of the histological subtype of non-small-cell lung cancer using MRI texture analysis. Biomed Signal Process Control 66:102446
6. Kumar JLM, Rashid M, Musa RM, Razman MAM, Sulaiman N, Jailani R, Abdul Majeed APP (2021) The classification of EEG-based wink signals: A CWT-Transfer learning pipeline. ICT Express. https://doi.org/10.1016/j.icte.2021.01.004
7. Mahendra Kumar JL, Rashid M, Muazu Musa R, Mohd Razman MA, Sulaiman N, Jailani R, Abdul Majeed APP (2021) The classification of EEG-based winking signals: a transfer learning and random forest pipeline. Peer J 9:e11182 . https://doi.org/10.7717/peerj.11182
8. Hany M (2020) Chest CT-Scan images dataset. https://www.kaggle.com/datasets/mohamedhanyyy/chest-ctscan-images

9. Abdullah MA, Ibrahim MAR, Shapiee MNA, Zakaria MA, Razman MAM, Musa RM, Osman NAA, Majeed APPA (2021) The classification of skateboarding tricks via transfer learning pipelines. Peer J Comput Sci 7:e680
10. Noor FNM, Mohd Isa WH, Khairuddin IM, Razman MAM, Jizat JAM, Nasir AFA, Musa RM, P. P. Abdul Majeed A (2021) The diagnosis of diabetic retinopathy: a transfer learning with support vector machine approach. Adv Intell Syst Comput 1350 AISC:391–398

Chapter 5
Skin Cancer Diagnostics: A VGGEnsemble Approach

Abstract The human skin is the largest organ of the human body, and it is highly susceptible to lesions. This study attempts to classify two distinct classes of malignant skin cancers, i.e., Actinic Keratosis (AK) and Basal Cell Carcinoma (BCC), as well as Dermatofibroma (DF), which is benign. A total of 330 dermoscopy images were split into the 70:15:15 ratio for training, testing and validation, respectively. Different VGG-Logistic Regression (LR) pipelines, i.e., VGG16-LR and VGG19-LR, were formulated. In addition, the effect of combining the features extracted from both VGG models, dubbed as VGGEnsemble, was also investigated. It was demonstrated from the study that the ensemble model yielded a better classification accuracy than its standalone versions. Therefore, it could be concluded that the performance of the pipeline is improved through this approach and subsequently could aid the diagnostics of different types of skin diseases by dermatologists.

Keywords Computer-Aided diagnosis · Skin cancer · VGG16 · VGG19 · Ensemble

5.1 Introduction

As the largest organ of the human body, the skin is susceptible to lesions. Skin cancer has been on the rise across the globe over the past decade, and WHO suggests that one-third of cancers diagnosed are skin cancer [1, 2]. It is amongst the rapid-growing fatal cancer. However, early detection of such cancer allows for treatment to be sought and subsequently, increases the chance of living. The diagnosis of skin lesions is often carried out by dermatologists via dermoscopy assessment; nevertheless, such a technique is somewhat prone to misdiagnosis owing to the complex nature of the lesions that at times are indiscernible. With the advent of rapid computational technology and artificial intelligence, the aforesaid predicament could be mitigated through computer-aided diagnostics (CAD).

M. H. Arzmi et al., *Deep Learning in Cancer Diagnostics*,
SpringerBriefs in Forensic and Medical Bioinformatics,
https://doi.org/10.1007/978-981-19-8937-7_5

Hitherto, significant advances have been made by researchers in the diagnosis of skin lesions via both machine learning and deep learning approaches. Hassan et al. [3] developed a vanilla CNN model for the classification of benign and malignant skin lesions. The dermoscopic images were extracted from the International Skin Imaging Collaboration (ISIC) archive. They investigated the effect of the number of hidden neurons and the number of iterations towards the classification accuracy of the classes. It was reported that the pipeline developed is able to achieve a recall of 0.84.

Ali et al. [4] investigated the efficacy of different pre-trained Convolutional Neural Network (CNN) models reported in the literature, namely AlexNet, ResNet, VGG16, DenseNet and MobileNet, against their proposed deep CNN (DCNN) models in classifying benign and malignant skin lesions respectively based on the HAM10000 dataset. Different hold-out cross-validation techniques 80:10:10 and 70:20:10 for training, validation and testing, respectively. Moreover, it is worth noting that different pre-processing techniques were also employed prior to the development of the DCNN. It was shown that their proposed model is able to classify skin lesions well.

Wang et al. [5] compared the performance of a pre-trained CNN model, i.e., ResNet, against conventional machine learning models, namely Support Vector Machine (SVM), k-Nearest Neighbour (kNN) and Random Forest (RF) to classify benign and malignant classes. The fivefold cross-validation technique was employed. The Principal Component Analysis (PCA) was used to extract the features from the images. It was demonstrated through the study that the ResNet model could achieve an average accuracy of 78%, whilst the other evaluated models yielded an average accuracy of 60%.

The ability of combined CNN (cCNN) in the diagnosis of non-pigmented skin cancer was carried out by Tschandl et al. [6]. The dataset consists of dermoscopic as well as close-up images of nine types of skin lesions. Different data augmentation techniques were performed on the image input. It was shown from the study that the InceptionV3-based architecture could distinguish the dermoscopic images better, whilst the ResNet50 is better in discerning clinical close-up images. It was reported that the integration of the architectures provided a better classification of the classes generally with respect to the types of input images. Moreover, the study further demonstrated that the cCNN architecture classifies non-pigmented lesions as good as expert raters and much better than less-experienced raters suggesting that AI-based CAD could facilitate the diagnosis of complex skin lesions.

It is apparent from the aforementioned literature that the employment of deep learning has made a significant contribution towards skin cancer-based CAD. Nonetheless, there is an apparent gap in the ensemble of the features extracted via pre-trained CNN models and replacing the fully connected layers with a conventional machine learning model. Therefore, in the present study, the effect of concatenating the features extracted from the pre-trained CNN models, particularly the VGG family, viz. VGG19 and VGG16 towards discerning different types of benign and malignant tumours, i.e., Actinic Keratosis (AK), Basal Cell Carcinoma (BCC) and Dermatofibroma (DF) are investigated.

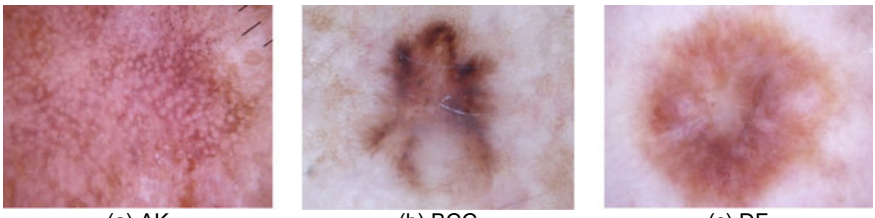


Fig. 5.1 Types of Skin Lesions (**a**) AK (**b**) BCC and (**c**) DF

5.2 Methodology

The dataset was retrieved from an open-source repository, i.e., Kaggle [7] sourced from the International Skin Imaging Collaboration (ISIC). A total of 330 dermoscopy-based images of AK, BCC and DF is used in the present study from the repository. The aforesaid classes are depicted in Fig. 1a, c. It is worth noting that undersampling was performed in order to have an equal representation of the types of skin lesions to mitigate biased representation of the data. The dataset was split into the 70:15:15 ratio for training, testing and validation, respectively. In the present study, the VGG family, i.e., VGG16 and VGG19, were employed to extract the features from the CT-based images. The proposed architecture in the present investigation is a departure from the traditional means of deploying a Convolutional Neural Network (CNN) model, i.e., the fully connected layers are swapped with a conventional machine learning model [8–10]. In addition, an ensemble between VGG16 and VGG19 is formed by concatenating the features [11, 12], dubbed as VGG16Ensemble. The features were then fed into a Logistic Regression (LR) classifier. It is worth mentioning that the hyperparameters are taken as default from the scikit-learn library. The analysis was carried out using a Python IDE, namely Spyder 3.3.6, along with its associated Keras and TensorFlow libraries. The performance of the pipelines is evaluated via the classification accuracies, as well as the macro average of the precision, recall and f1-scores, respectively.

5.3 Results and Discussion

Figure 5.2 illustrates the performance of the investigated pipelines. It is evident that generally, the VGGEnsemble-LR pipeline yield a better accuracy with respect to the train, validation and test dataset against the standalone feature extracted pipelines. A similar increase in performance has been reported through the use of the ensemble technique [12]; this is due to the leveraging of the features contributed by both models.

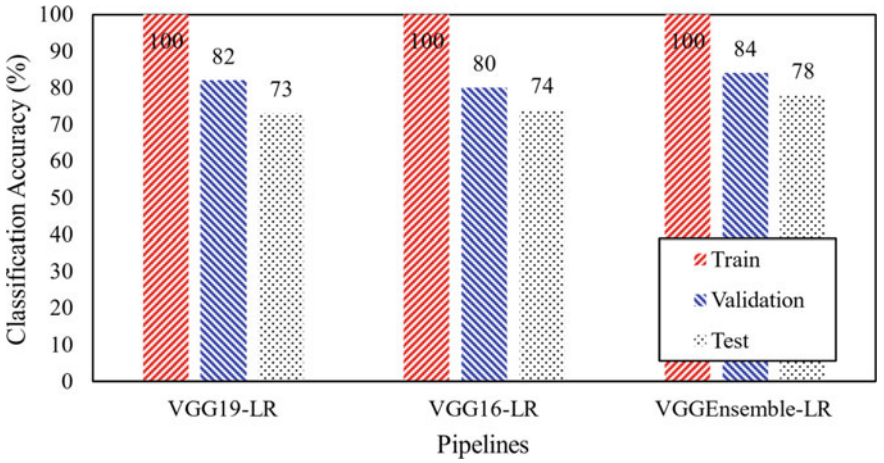


Fig. 5.2 The performance of the different VGG-LR pipelines evaluated in terms of classification accuracy

Table 5.1 lists the macro average of the precision, recall and f1-scores, respectively, for the pipelines appraised from the validation dataset. It is apparent from the performance measures that the VGGEnsemble-LR pipeline is significantly better compared to the other evaluated pipelines. Figure 5.3 depicts the confusion matrix on the test dataset of the VGGEnsemble-LR pipeline. The classes AK, BCC and DF are denoted as 0, 1 and 2, respectively. It could be seen that there are several misclassifications that transpired. Nevertheless, it is worth noting that there is only a 6% chance that malignant lesions such as AK and BCC are misclassified as DF, a benign lesion. Conversely, four benign cases were misclassified as malignant cases, which is somewhat desirable in some sense, suggesting that the model is more conservative than the. The misdiagnosis could arise from the indistinguishable traits between the type of lesions and the quality of the selected dermoscopic images.

Table 5.1 Other performance measures on the validation dataset

Pipeline	Performance Measures (%)		
	Precision	Recall	F1-Score
VGG16-LR	82	82	82
VGG19-LR	81	80	80
VGGEnsemble-LR	84	85	85

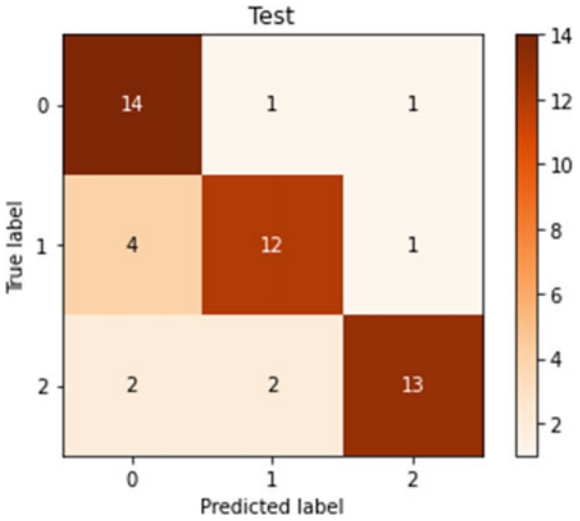


Fig. 5.3 The confusion matrix of the VGGEnsemble-LR pipeline on the testing dataset

5.4 Conclusion

In this chapter, the effect of concatenating the features extracted from pre-trained CNN models, viz. VGG16 and VGG19 towards the classification efficacy of skin lesions was demonstrated. It is evident from the study that the classification performance of the ensembled features is far superior to that of extracted standalone features. The proposed approach could be further improved by optimising the hyperparameters of the classifier or by replacing it with different types of classifiers. Nonetheless, it could be deduced that such a method is promising and could further facilitate dermatologists in discerning the different types of skin lesions, especially malignant ones.

References

1. Naji S, Jalab HA, Kareem SA (2019) A survey on skin detection in colored images. Artif Intell Rev 52:1041–1087. https://doi.org/10.1007/S10462-018-9664-9
2. Pacheco AGC, Krohling RA (2019) Recent advances in deep learning applied to skin cancer detection
3. Hasan M, Barman S das, Islam S, Reza AW (2019) Skin cancer detection using convolutional neural network. ACM International Conference Proceeding Series 254–258. https://doi.org/10.1145/3330482.3330525
4. Ali MS, Miah MS, Haque J, Rahman MM, Islam MK (2021) An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models. Machine Learning with Applications 5:100036. https://doi.org/10.1016/J.MLWA.2021.100036
5. Wang Y, Louie DC, Cai J, Tchvialeva L, Lui H, Jane Wang Z, Lee TK (2021) Deep learning enhances polarisation speckle for in vivo skin cancer detection. Opt Laser Technol 140:107006. https://doi.org/10.1016/J.OPTLASTEC.2021.107006

6. Tschandl P, Rosendahl C, Akay BN, Argenziano G, Blum A, Braun RP, Cabo H, Gourhant JY, Kreusch J, Lallas A, Lapins J, Marghoob A, Menzies S, Neuber NM, Paoli J, Rabinovitz HS, Rinner C, Scope A, Soyer HP, Sinz C, Thomas L, Zalaudek I, Kittler H (2019) Expert-Level Diagnosis of Nonpigmented Skin Cancer by Combined Convolutional Neural Networks. JAMA Dermatol 155:58–65. https://doi.org/10.1001/JAMADERMATOL.2018.4378
7. Skin Cancer ISIC | Kaggle. https://www.kaggle.com/datasets/nodoubttome/skin-cancer9-classesisic. Accessed 14 Oct 2022
8. Abdullah MA, Ibrahim MAR, Shapiee MNA, Zakaria MA, Razman MAM, Musa RM, Osman NAA, Majeed APPA (2021) The classification of skateboarding tricks via transfer learning pipelines. PeerJ Comput Sci 7:e680
9. Mahendra Kumar JL, Rashid M, Musa RM, Mohd Razman MA, Sulaiman N, Jailani R, P.P. Abdul Majeed A (2021) The classification of EEG-based wink signals: A CWT-Transfer Learning pipeline. ICT Express. https://doi.org/10.1016/j.icte.2021.01.004
10. Mahendra Kumar JL, Rashid M, Muazu Musa R, Mohd Razman MA, Sulaiman N, Jailani R, Abdul Majeed APP (2021) The classification of EEG-based winking signals: a transfer learning and random forest pipeline. PeerJ 9:e11182. https://doi.org/10.7717/peerj.11182
11. Almanifi ORA, Razman MAM, Khairuddin IM, Abdullah MA, Majeed APPA (2021) Automated Gastrointestinal Tract Classification Via Deep Learning and the Ensemble Method. In: International Conference on Control, Automation and Systems. pp 602–606
12. Almanifi ORA, Mohd Khairuddin I, Mohd Razman MA, Musa RM, P.P. Abdul Majeed A, (2022) Human activity recognition based on wrist PPG via the ensemble method. ICT Express. https://doi.org/10.1016/j.icte.2022.03.006

Chapter 6
The Way Forward

Abstract This chapter shall summarise the different approaches of feature-based transfer learning that have been employed in the book as well as provides some future direction worth scrutinising with regards to AI-driven CAD of cancers specifically and medical imaging in general.

Keywords Transfer learning · Computer-Aided-Diagnostics · Deep learning · Data-Centric AI

6.1 Summary

In this book, the authors have examined different feature-based transfer learning architectures towards facilitating the shift towards AI-driven Computer-Aided-Diagnostics (CAD) of different types of cancer, namely lung, breast, skin and oral, respectively, in discerning its malignancy. It is important to note at this juncture, before dissecting the evaluated architectures, the discussion ought to be focused on the data preparation itself. The quality of data is rather crucial in order to develop a robust pipeline, and having an unbiased representation of the data is non-trivial. There are instances in the literature in which data skewness transpired and, ultimately begs the question of its actual robustness, albeit excellent classification accuracies are reported. In the same vein, the authors echo Andrew Ng on the need for a data-centric approach [1], even for medical imaging diagnostics, as more is not necessarily good; nonetheless, a careful selection of quality data could do wonders.

It has been demonstrated in the book that different feature-based transfer learning architectures could possibly yield a desirable classification of the different types of cancers. To reiterate, this book is dedicated to an approach in which the features are extracted from pre-trained CNN models with its fully connected layer swapped with a conventional machine learning model. The merits of such an approach have been discussed in [2–5]. Amongst the architectures that have been discussed in this book is by exploring the performance of different transfer learning models (for feature extraction) with one specific machine learning model. This somewhat indicates that a particular transfer learning model has the ability to extract better features from a given

M. H. Arzmi et al., *Deep Learning in Cancer Diagnostics*,
SpringerBriefs in Forensic and Medical Bioinformatics,
https://doi.org/10.1007/978-981-19-8937-7_6

image from another. Taking a cue from that, the effect of combining (concatenating) the features from different transfer learning models was also subsequently investigated, and it was established from the analysis carried out that a better diagnostic was achieved. The effect of the selection of different classifiers with one transfer learning model was also examined, suggesting that different classifiers would have an edge against another based on the type of image or specifically the features that were extracted from a specific transfer learning model. In addition, optimising the hyperparameters of a given model has also been demonstrated to improve a pipeline's performance.

The present work did not investigate the effect of exploiting different image preprocessing techniques prior to the feature extraction phase. This was deliberately done to illustrate the natural capability of the transfer learning pipelines themselves. Therefore, it is given that some of the analysis illustrated in the book does not indicate excellent classification accuracies. However, it does conceptualise the nuances of the overall framework that could be emulated in different medical imaging applications that pave the way for a more robust CAD to facilitate medical practitioners upon its deployment.

References

1. Andrew Ng: Unbiggen AI—IEEE spectrum. https://spectrum.ieee.org/andrew-ng-data-centric-ai. Accessed 15 Oct 2022
2. Mahendra Kumar JL, Rashid M, Musa RM, Mohd Razman MA, Sulaiman N, Jailani R, Abdul Majeed APP (2021) The classification of EEG-based wink signals: A CWT-transfer learning pipeline. ICT Express. https://doi.org/10.1016/j.icte.2021.01.004
3. Mahendra Kumar JL, Rashid M, Muazu Musa R, Mohd Razman MA, Sulaiman N, Jailani R, Abdul Majeed APP (2021) The classification of EEG-based winking signals: a transfer learning and random forest pipeline. Peer J 9:e11182. https://doi.org/10.7717/peerj.11182
4. Abdul Rauf AR, Mohd Isa WH, Khairuddin IM, Mohd Razman MA, Arzmi MH, Abdul Majeed APP (2022) The classification of oral squamous cell carcinoma (OSCC) by means of transfer learning. Lect Notes Networks Syst 429 LNNS:386–391. https://doi.org/10.1007/978-3-030-97672-9_34/FIGURES/5
5. Noor FNM, Majeed APPA, Razmam MAM, Khairuddin IM, Isa WHM (2021) The diagnosis of diabetic retinopathy: a transfer learning approach. In: International Conference on Control, Automation and Systems. pp 596–601