We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Chapter

Approaches for Handling Immunopathological and Clinical Data Using Deep Learning Methodology: Multiplex IHC/IF Data as a Paradigm

Siting Goh, Yueda Chua, Justina Lee, Joe Yeong and Yiyu Cai

Abstract

Recent advancements in deep learning based artificial intelligence have enabled us to analyse complex data in order to provide patients with improved cancer prognosis, which is an important goal in precision health medicine. In this chapter, we would be discussing how deep learning could be applied to clinical data and immunopathological images to accurately determine survival rate prediction for patients. Multiplex immunohistochemistry/immunofluorescence (mIHC/IF) is a relatively new technology for simultaneous detection of multiple specific proteins from a single tissue section. To adopt deep learning, we collected and pre-processed the clinical and mIHC/IF data from a group of patients into three branches of data. These data were subsequently used to train and validate a neural network. The specific process and our recommendations will be further discussed in this chapter. We believe that our work will help the community to better handle their data for AI implementation while improving its performance and accuracy.

Keywords: immunopathology, deep learning, multiplex IHC/IF

1. Introduction

Improved cancer prognosis is a vital goal of precision health medicine. Advancements in Deep Learning (DL) based Artificial Intelligence (AI) technologies enable modelling of complex data providing deeper insights and patients with more reliable results. Machine Learning (ML) is the process of enabling machines to make predictions from data that is fed into it. This includes DL, a type of approach created from the development of artificial neural networks [1]. The DL network consists of multiple layers of artificial neural networks including an input, an output and multiple hidden layers [2, 3]. Predictions are made after datasets are generated from and trained against these hidden layers. Recent advancements in computational processing power has sparked interest in tapping into the vast research on DL and applying it to digital pathology. Digital pathology is the process of digitizing Whole-Slide Images (WSI) using advanced slide-scanning techniques and AI-based methods for detecting, segmenting, diagnosing and analysing digitized images [4].

DL is the engine of advancement in artificial learning in both computer and clinical sciences. It is a collection of computer learning algorithm layers that uses the raw data input to first generate generalise features that are subsequently used to progressively extract higher level features such as tumour stroma count, and assign them class labels. Eventually, the system will distinguish different interest categories via the identified ideal data features. The DL approaches are widely accepted due to the ability of discovering patterns and signals from data too large for human comprehension. Furthermore, the multiple layers allow modelling of highly complex non-linear problems. On top of having higher accuracy, the DL approaches are also easily applied.

1.1 The importance of deep learning in digital pathology and mIHC

In current clinical practice, pathologists base their medical diagnosis on the quantification and visual recognition of the analysed sample details, which could lead to diagnostic discrepancies and potential suboptimal patient care [5]. The increased adoption of non-invasive clinical procedures to acquire diagnostic samples has also severely reduced the quantity and quality of samples obtained, which compounds the workload of pathologists. In view of the inter-variabilities in analysing samples manually and the limitations of available samples, the use of DL analysis has thus been researched on and progressively applied in the clinical practice.

DL in digital pathology aims to improve the workload of pathologists by automating time-consuming tasks, hence allowing additional time to be spent on disease presentations with complex features. AI applications in digital pathology can also be applied to develop prognostic assays that evaluate the severity of diseases and make an accurate prognosis in response to therapy. This could be applied to various image processing and classification tasks, such as low-level jobs revolving around image recognition issues including detection and segmentation, as well as high-level tasks such as prognosis of response to therapy based on patterns of images [6, 7]. Such AI approaches are designed to extract relevant image renditions to train machines to be used as specific segmentation, diagnostic or prognostic tools.

One of the most extensively used DL models in pathology image analysis is the Convolutional Neural Networks (CNN). The CNN is a class of deep, feedforward networks, comprising several layers which extrapolate an output from an input and contains multiple convolutional sheets. These convolution sheets are the foundation of a CNN in which the network learns and extrapolates feature maps from images using filters between the input and output layers [4]. These layers in CNN are not connected as the neurons in one layer only interact with a specific region of the previous layer instead of all its neurons. The CNN also contains pooling layers which primarily function to scale down or reduce the dimensionalities of the features. CNN DL-based approaches are used for image-based detection and segmentation tasks to distinguish and quantify cells, histological features or highlight regions of interest [4]. CNN DL-based approaches have also been developed to automatically distinguish and segment blurry areas in digitised WSIs with high accuracy.

Another type of DL approach is the Fully Convolutional Network (FCN) which learns representations from every pixel and makes a potential feature detection that may occur sparsely in the entire pathology image [4]. FCNs uses co-registered Haemotoxylin and Eosin (H&E) images with multimodal microscopy techniques to classify WSIs into 4 classes: cancer, non-malignant epithelium, background and other tissues. When FCN was used to detect invasive breast cancer regions on WSIs, it had a diagnostic accuracy of 71% (Sørensen-Dice coefficient) when compared to an expert breast pathologist's assessment [8]. With better technologies and further research, FCNs can potentially automate these tasks with a higher accuracy, reducing the workload of pathologists.

AI-based approaches such as Generative Adversarial Network (GAN)-based approaches can be used for training to automatically score tumoral programmed cell death 1 ligand 1 (PD-L1) expression in biopsy sample images [4]. They reduce the number of inputs required from pathologists and make up for lack of tissue samples available in biopsy specimens. Novel GAN-based approaches propose converting H&E staining of WSIs to virtual immunohistochemistry staining, thus eliminating the need for destructive IHC tissue testing.

Many have also trialled DL in the field of immunohistochemistry (IHC). Traditional IHC is a common diagnostic tool used in pathology, but its application is significantly limited by its ability of only allowing single marker labelling per tissue section [9]. Alternatively, multiplex immunohistochemistry/immunofluorescence (mIHC/IF) technologies permit simultaneous detection of several markers on a single tissue section [9]. However, analysing large samples with multiple markers in conventional and manual ways by pathologists are highly time-consuming, laborious and susceptible to human error. By combining mIHC/IF with DL to analyse digitized WSIs, this will overcome the limitations.

In conclusion, the research and diagnostic fields have come a long way since the introduction of IHC. With the introduction of Al-based approaches in the application of IHC, higher accuracy and productivity could be achieved not just in the diagnostic level but also providing us with a platform to further venture into areas of medical knowledge yet to be fully understood.

1.2 mIHC/IF Technologies

To-date, our understanding of cancer immunotherapy has evolved and led to multiple studies investigating and refining strategies targeting negative regulators. Many have studied the use of checkpoint blockade immunotherapy such as programmed cell death receptor 1 (PD-1), PD-L1 and cytotoxic T-lymphocyte–associated protein 4 (CTLA-4) in a variety of cancers. The subsequent success of checkpoint blockade inhibition in clinical trials has led to the Food and Drug Administration's approval of various drugs such as Ipillimumab and Prembrolizumab for melanoma treatment of non-small cell lung cancer (NSCLC) respectively [10]. Furthermore, trial of combination immunotherapy has shown clinical efficacy in various cancers [11, 12]. However, other studies have also suggested that efficacy of these immunotherapy in various cancers may depend on the expression of biomarkers. For example, PD-L1 is suggested as a useful predictive marker in patients with NSCLC receiving Prembrolizumab [13]. However, this is not the case in patients with stage III melanoma [14]. To further discover potential biomarkers that could determine the efficacy of immunotherapy in various cancers, IHC has been introduced as a platform for these clinical studies.

Since its introduction in the 1940s [15], conventional IHC has been widely used in field of pathology and research. It involves the process of staining tissues samples using antibodies specific to antigens present within the samples. This specificity allows microscopic visualisation for diagnosis of neoplasm and obtaining valuable prognostic information. Despite this, it does have several limitations. The inability of labelling more than one marker per tissue sample has resulted in loss of potential information for analysis. For instance, the prediction of prognosis to an immunotherapy such as PD-L1/PD-1 checkpoint blockade may depend on the expression of an individual biomarker or in combination with other biomarkers [16–18]. Furthermore, the immune system can potentially be better understood, if the analysis of various biomarkers' expression patterns are done simultaneously, or cellular interactions within the tumour microenvironment can be visualised [19].

Moreover, IHC involves many critical steps which have high inter-user variability. For instance, antigens such as Ki-67 are more vulnerable to ischemia. As such, over fixation could result in irreversible damage to these antigens [20]. The concern of IHC's reproducibility such as for Ki-67 and its implications was also mentioned in the 2017 St. Gallen International Expert Consensus Conference [21]. However, multiple studies have since demonstrated that analytical variability can be negated with the use of digital analysis to calculate biomarkers index [22, 23].

Although conventional IHC is a cost-effective diagnostic and prognostic tool, it has been replaced with the introduction of mIHC. mIHC has been used to overcome the shortcoming of single biomarker labelling in conventional IHC. The use of mIHC has proven to provide an even more accurate analysis as seen in the study by Yeong *et al.,* where the simultaneous quantification of three different PD-L1 antibodies (22C3, SP142 and SP263) by mIHC scoring had moderate-to-strong correlation (with 67%–100% individual concordance rates and Spearman's rank correlation coefficient values up to 0.88 [24]) when compared with manual scoring from four different pathologists.. This demonstrated the use of mIHC as a promising tool for an even more accurate analysis.

The use of mIHC has played a significant role in both research and clinical studies of cancer immunotherapy. mIHC is a relatively new tool to study the spatial tumour microenvironment especially those of limited tissue specimens. It has great potential in clinical and translational application. This was demonstrated by Halse *et al.*, who used mIHC to reveal a close relationship between the presence of $CD8^+$ T cells within the tumour and the expression of PD-L1 in melanoma [25]. A systemic review and meta-analysis of studies also reported that mIHC improved results in predicting responses to PD-1/PD-L1 checkpoint blockade immunotherapy in various solid tumour types when compared to using conventional IHC analysis [26]. Several studies have also used various types of mIHC to obtain data for analysis. For instance, TSA-based mIHC was used to profile PD-1 to PD-L1 proximity in 166 metastatic melanoma samples and 42 Merkel cell carcinoma samples in two respective studies [27, 28]. As aforementioned, understanding the tumour microenvironment could potentially provide a foundation upon which interpretation of immunotherapy response could be made.

1.3 Use of mIHC in combination with digital pathology

mIHC can be powered by digital pathology analysis software, such as inForm (Akoya Biosciences, California, USA) [29–31] and HALO TM (Indica Labs) [28, 32].

	Digital pathology software		
	InForm	**HALO**	**Oncotopix**
Developer	**Akoya Biosciences**	**Indica Labs**	**Visiopharm**
Compatibility with multiplex IHC platforms	Yes	Yes	Yes
Co-localisation of markers	Yes	Yes	Yes
Tissue segmentation	Yes	Yes	Yes
Spatial analysis	No	Yes	Yes
Ready solution for interrogation of breast cancer markers	No	No	Yes (Ki-67, HER2, ER, PR)
Use in breast cancer research	Yes	Yes	Yes

Table 1.
Digital pathology softwares, InForm, HALO, and Oncotopix and their software features for multiplex IHC/IF.

These software resolve the restrictions of labeling a single marker per tissue section by precisely evaluating the unique localization of multiple simultaneously detected biomarkers and their co-expressions or interactions between cells [33].

For example, although Ki-67/PD-L1 labeling is useful by itself, a multiplex approach enables several markers to be interrogated simultaneously [34–36]. However, only analytical digital pathology solutions for Ki67 and PD-L1 scoring are currently commercially available as listed in **Table 1** [33, 37]. The involvement of digital pathology has also decreased intra- and inter-observer variability seen in manual scoring as previously highlighted. Consequently, using mIHC in conjunction with digital analysis software will resolve the restrictions of conventional IHC, thus providing us with an accurate and powerful tool in the interpretation of immune response in various fields.

2. Proposed deep learning framework for analysing immunopathological and clinical data

This section presents a holistic guiding framework to select and develop a DL architecture for multi-dimensional analysis. The entire pipeline can be broken down into 3 parts: [1] data pre-processing, [2] feature engineering and [3] model selection, validation, and evaluation (**Figure 1**). This includes treating the data input, selecting the appropriate model for the type of data and using the preferable method to validate the selected model.

To demonstrate the clinical application of the framework, a total of 107 clinical as well as mIHC/IF data from patients with breast cancer (BC) previously published [37]. The clinical data consists of parameters such as age and tumour grade as stated in **Table 2** Row 1, while the mIHC data comprised of antibody-based spectral unmixing result obtained from stained mIHC image of tumour section labeling markers such as cytokeratin, CD68, CD8, CD20, FOXP3, PD-L1, and CK/EpCAM (**Figure 2**).

2.1 Data pre-processing

The first step in data pre-processing involves analysis of the dataset. This process consists of four main components: [1] one-hot encoding, [2] data normalization, [3] data enhancement and [4] data shape conformity.

2.1.1 One-hot encoding

One-hot encoding is the process of converting any non-numerical data existing in the clinical dataset to a categorical numerical representation that is readable by

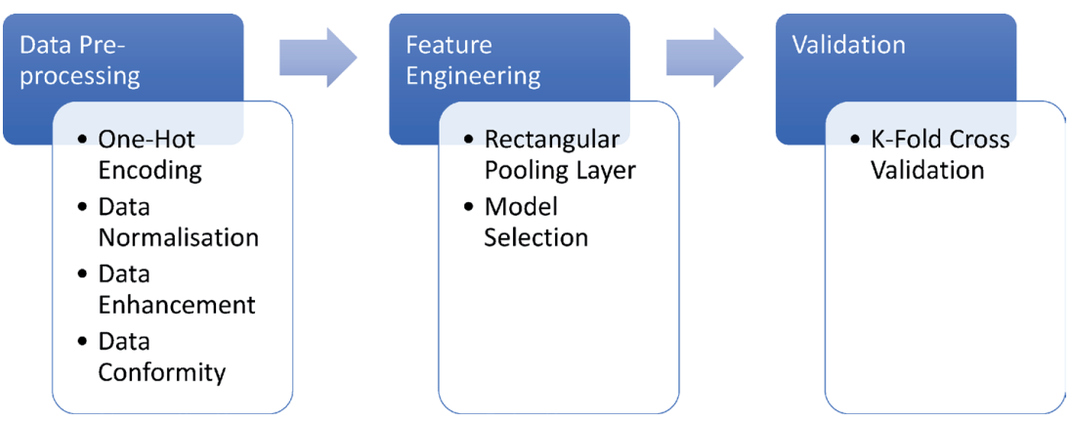


Figure 1.
General overview of the DL framework.

Age at Diagnosis	Tumour Grade	Tumour Size	Lymphovascular Invasion	Lymph Node Positive	Lymph Node Stage	Disease Free (month)	Overall Survival (month)
Integer	Ordinal Integer	Integer	Categorical (positive/ possible/negative)	Integer	Ordinal Integer	Integer	Integer

Table 2.
Data of Clinical Dataset.

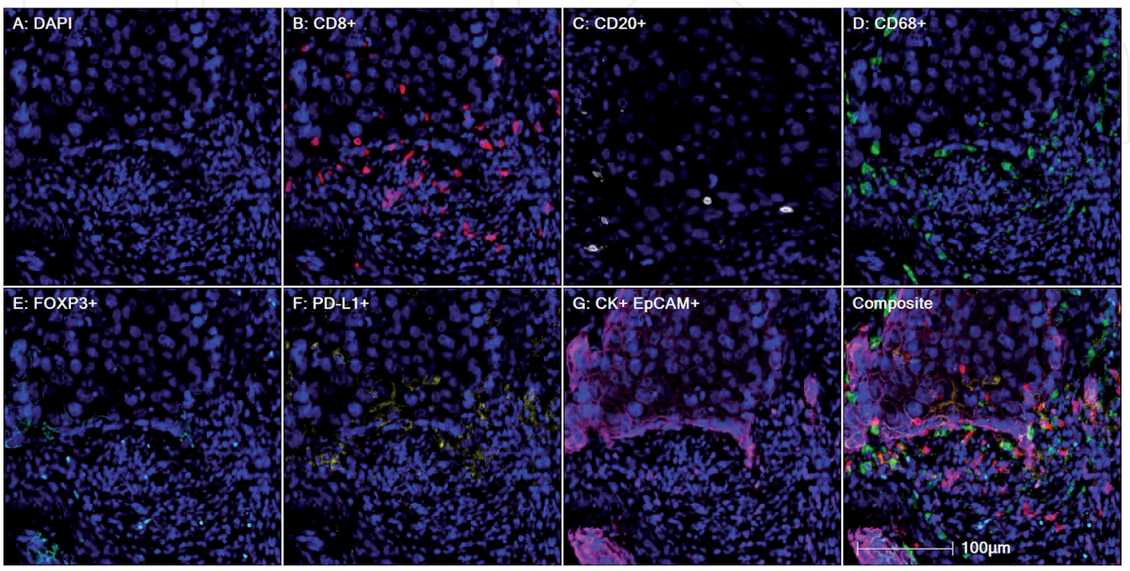


Figure 2.
Representative images of breast tissue stained using multiplex immunohistochemistry/immunofluorescence (mIHC/IF) [DAPI (blue), CD8 (red), CD20 (white), CD68 (green), FOXP3 (cyan), PD-L1 (yellow), CK/EpCAM (magenta)]. (Magnification, 200X).

the computer. Any non-numerical data within each category is split into the number of categories it has and encoded with a binary 0/1. For example, in the case of our clinical data, the columns, "Lymphovascular Invasion" contains 3 possible values: positive, possible, negative (**Table 3**). This represents 3 categories and is the prime candidate to be one-hot encoded. The input column is subsequently expended to 3 columns, one for each category in this input as shown in **Tables 3** and **4**.

2.1.2 Data normalisation

Most CNN research and models are developed with the intention for application in Computer Vision, where an entire input image data points are all pixels with ranges from zero to 255. Non-imaging datasets are more complicated as each input parameters have different units of measurements that might range from ones

Lymphovascular Invasion
Absent
Present
Absent
Absent
Possible

Table 3.
Lymphovascular invasion data before One-hot encoding.

Lymphovascular Invasion Absent	Lymphovascular Invasion Possible	Lymphovascular Invasion Present
1	0	0
0	0	**1**
1	0	0
1	0	0
0	**1**	0

Table 4.
Lymphovascular invasion data after One-hot encoding.

to hundreds of thousands. Using such models with disparate values meant that a model with a large input parameter could easily outweigh another with a smaller value range. Therefore, data normalisation is needed to ensure that the dataset has comparable values across the data inputs while still maintaining their distribution within each data input. Data normalisation was done by scaling each input column to carry a mean of 0 and a standard deviation, by applying the following formula:

$$X_{New} = \frac{X - \mu}{\sigma}$$

An example of a segment of clinical dataset following one-hot encoding is as shown in **Table 5**. A notable feature of the clinical dataset is the disparate values across the columns which arose due to the different units of measurements used across the columns, such as categorical numbers, months, and millimetres. As such, these numbers could not be directly compared. To obtain a more comparable data, normalisation of these values was done, while maintaining the distribution within each column (**Table 6**).

2.1.3 Data enhancements

When working with a medical dataset, it will be advantageous to have medical insights augment the data, as it can improve the result. The use of medical insights is however dependent on the context of the problem and is subjective to the augmentation or removal of features and/or any dataset. In this study, clinically relevant data was augmented to the cell dataset to count the number of stroma and cancer cells of each patient. Subsequently this was evaluated with a simple 12-layer dense neural network and the obtained results were compared with and without data enhancements on 10000 epochs. It was discovered that there was a marked improvement in the reduction of mean absolute error by 14.8% when the clinical dataset was enhanced with more relevant information. However, the reduction in mean absolute error was highly dependent on clinical dataset used and thus varies with its application.

2.1.4 Ensuring data conformity

The CNN requires the dataset to be homogeneous in its shape, which is achievable in the classical Computer Vision problems where images could be resized to a uniform rectangular shape. However, in the case of medical dataset, the dimensions are mostly dependent on the source of the data, which is usually 3-dimensional or more. There are two ways to homogenise medical dataset, either by appending non-meaningful data to the clinical dataset, or selectively removed data until the shape is uniform. This process requires a higher dimensional visualisation which is best explained using a tangible example as follows:

Age at Diagnosis	Tumour Grade	Tumour Size (mm)	Lympho-vascular Invasion (Present)	Lymph Node Positive	Lymph Node Stage	tumour count	stroma count	Disease Free Survival (month)	Overall Survival (month)
54	3	25	0	0	0	1550	667	39.90322581	48.5483871
43	3	30	1	0	0	4236	1144	27.5483871	32.12903226
47	3	30	0	1	1	3132	1715	140.8666667	140.8666667
36	3	50	0	0	0	6264	777	139.3225806	139.3225806
49	3	24	1	0	0	3404	1504	139.1	139.1
65	2	32	0	0	0	3338	5237	50.6	139.2666667

Table 5.
Original Sample Cell Data Before Normalisation.

Age at Diagnosis	Tumour Grade	Tumour Size (mm)	Lymphovascular Invasion (Present)	Lymph Node Positive	Lymph Node Stage	tumour count	stroma count	Disease Free Survival (month)	Overall Survival (month)
0.01679224	0.35540932	0.32887951	0.69721667	0.09712858	0.37974281	0.65319726	0.83253395	0.85909418	0.75847927
0.91517705	0.35540932	0.15397838	1.4342743	0.09712858	0.37974281	0.65319726	0.93975876	0.37551795	1.00897651
0.58849167	0.35540932	0.15397838	0.69721667	0.09712858	0.21114331	0.43886691	0.21131083	0.20335423	1.28857808
1.48687648	0.35540932	0.54562614	0.69721667	0.09712858	0.37974281	0.65319726	2.27788594	0.74757765	1.25727138
0.42514897	0.35540932	0.36385973	1.4342743	0.09712858	0.37974281	0.65319726	0.39078351	0.01055475	1.25275851
0.88159257	2.81365708	0.08401793	0.69721667	0.09712858	0.37974281	0.65319726	0.34723499	3.77391084	0.54159967

Table 6.
Sample Cell Data After Normalisation.

In this study, the cell dataset has the attributes of a 3D feature shape, where each patient has a list of cell inputs. This gives it a general 3D shape consisting of (Patients x No of Cells x Cell Parameters x 1). However, as each patient has a different number of cells, this results in a non-uniform dataset shape of (104 patients x χ cells x 144 Cell Parameters x 1), whereχ denotes the patient dependent number of cells in the dataset which ranges from 991 cell inputs to 10720 cell inputs.

Two strategies were evaluated to assess for the best approach of ensuring data conformity. The first method involves keeping the model and parameters constant. A dataset (Complete Cell Data) of "ghost cells" with value 0 were appended to the data to ensure regular shape of (107 x 10720 x 144). Alternatively, a dataset (Random Sampled Cell Data) with cells that were sampled randomly are pegged to the patient with the least number of markers to create a shape of (107 x 991 x 10720) (**Table** 7). Random sampled cell data are also employed for the training and evaluation of the DL model in this study. Following evaluation, no difference in the accuracy between both sets of data were observed. However, it was found that the smaller dataset required a smaller DL network that requires less computational power.

2.2 Feature engineering

In retrospect, there is no definitive answer when selecting a DL model. However, there is a wide array of models including the basic dense layers or CNN that could be applied. In view of multi-dimensional datasets, CNN is the most versatile in its ability to accommodate multi-dimensionality and has a strong community of research & development from Academics to Corporations. Despite so, CNN may be unsuitable for a non-imaging problem, as most CNN research is based on imaging problems where many of its tools such as max pooling may only work for spatial data. However, if harnessed correctly, CNN offers a highly flexible and advanced architecture that works for many types of data.

To understand the limitations of CNN on non-imaging dataset, it is essential to understand the fundamental difference between a spatial and non-spatial data. In spatial data like an image, a data point in one position is highly related to its surrounding pixels. Whereas for purely numerical data, one data point may not be related to its surrounding data points. It could instead be related to another data that is located at a different position, or more succinctly, is position independent.

2.2.1 Model selection and parameters

Most CNN tools assume that data points are position dependent. In this study, we looked at the dataset at hand to select a suitable CNN model and to adapt a powerful CNN tool called *Pooling Layer* to the non-spatial data. To select a suitable model that best fits the dataset and problem at hand, one should consider the general dimensions of the dataset which dictates the type of CNN to be used as listed in **Table 8**. For models that involve interaction with the environment, agent-based

Name	Size [patients*cell entries*markers]	Description
Base Dataset	107 x **12347** x 144	Full patient cell data, where patients with lacking cell entries are appended with "ghost cells" to form a regular shaped dataset
Random Row	107 x **991** x 144	Patients cell entries randomly selected via algorithm based on the patient with the least cell entries

Table 7.
Overview of Cell Data Size of Both Approaches.

Dimension	Probable Features (In imaging terms)	Best CNN Type
1D - Vector	(sample x features)	—
2D – Time Series/ sequence	(samples x timestamp x feature)	1D CNN
3D – Image Data	(samples x height x width x channels)	2D CNN
4D – Video Data	(samples x frames x channels x height x width)	3D CNN

Table 8.
Overview of Different Data Dimension and Suitable CNN Type.

Problem Type	Output Node Configuration	Objective Function	Equation
Scalar Regression	1 Node - Sigmoid Function	Mean Squared Error/ Mean Absolute Error	$MAE = \frac{1}{n}\sum \lvert y - \hat{y} \rvert$
Binary Classification	1 Node - Sigmoid Function	Cross Entropy	$BCE = \frac{1}{N}\sum_{i=1}^{N} y_i * \log\left(p\left(y_i\right)\right) + \left(1 - y_i\right) * \log\left(1 - y_i\right)$
Multi-Class Classification	1 Node for each class - SoftMax	Cross Entropy	$CE = -\sum_{i=N}^{N} y_i \log\left(f\left(s_i\right)\right)$

Table 9.
Overview of Objective Function of Different Problems.

models may be used. Furthermore, one should consider if the problem is a prediction or classification problem and if additional correlational features are necessary. As for a complex problem, the use of a deeper model may be more appropriate. Nevertheless, using a deeper model could lead to additional problems that require new architecture to overcome. Lastly, the objective of the problem and if it is a scalar prediction or classification problem must also be deduced (**Table 9**).

In this study, 2D CNN is used as the dataset is 3D and image-like. The problem type is a prediction model thus a plain 2D CNN with no agent-based is used. Given that the problem is complex, a Wide Residual Neural Network (WRN) identity mapping may help with modelling its complexity. Lastly, as this is a scalar prediction problem, the model should end with 1 sigmoid function and mean absolute error (MAE) objective function (**Table 9**).

The model selection process must be carefully chosen as it dictates the basis of the model and its result. To illustrate, an early stage proof-of-concept of application of DL on this dataset, a categorical approach was taken, where the patients were split into categories based on their survival rate in years. In designing it this way, the aim was to apply categorisation as the objective function [38]. However, this approach introduced an unintended consequence, a fixed error of the range of each category that could not be rid of regardless of how accurate the model is. This was because of framing a scalar problem as a categorical problem. Even though the resulting model achieved an accuracy of 90% [38], it did not show the in-built error of the prediction that was hidden by the range of each category.

Pooling layers progressively achieve spatial invariance by reducing the resolution of the feature maps, which reduces the number of parameters and computation in the network. This presents one with the ability to create a much deeper network with limited computational cost and overfitting. In a pooling layer, a simple function could be applied. The two conventional functions available are [1] maximization function, which find the maximum value of the region as a representation, and [2] average pooling function that aims to find an average representation of the region, where p is the resultant value of the pooling operation (**Figure 3**).

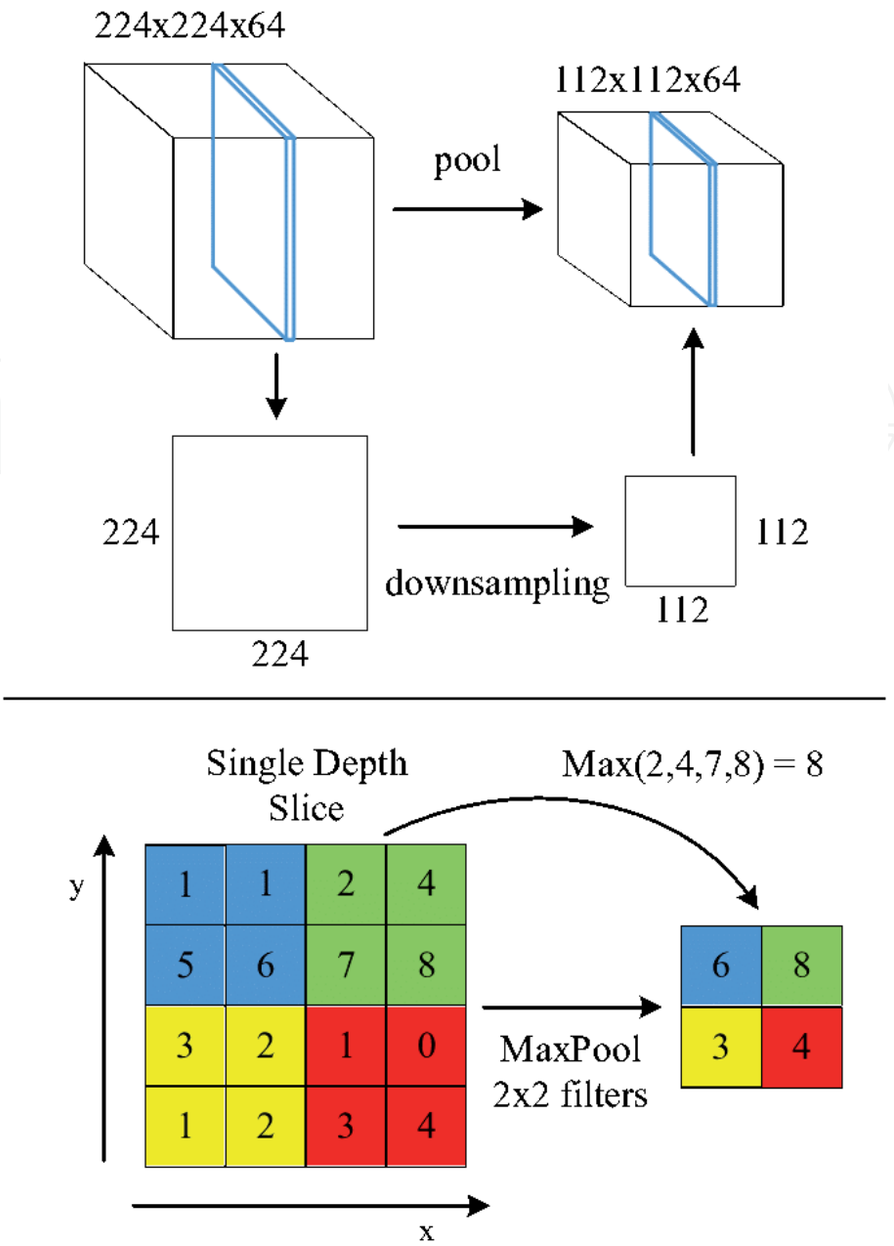


Figure 3.
Pooling Layer Computation & Representation- Pooling provides a form of abstraction of our data by down-sampling an input representation. There are two common rules for downsampling. Max-Pooling- which picks the input with the largest value. Average-Pooling – which averages out the input in the region. This prevents over-fitting by reducing noise in the data, also to reduce computational cost by reducing the number of parameters to learn. In the figure, a 4 x 4 matrix with 16 parameters is down-sampled to a 2 x 2 matrix of 4 parameters.

However, conventional form of a rectangular pooling layer is not applicable in datasets with only vertical relations. Pooling is mainly done in the context where all data in a 2D array are spatial, where any integers within the array size are related spatially. Such techniques help to compact representations, which could greatly influence the model's performance. With regards to this study, a 2D non-spatial cell dataset, each row has a different unit such as size (mm) or standard deviations. Pooling together variables of different types would result in an invalid representation. Thus, a different form of a pooling layer for non-spatial data could be created instead. Such rectangular pooling seeks to pool between data of the same type to create a representative value of the region, while reducing data noise, and the parameter size of the network.

Furthermore, the operation of the study aims to take a sample of group of nine cell markers of the BC dataset and to obtain the maximum value of each set. A graphical representation of the operation of max rectangular pooling layer (RPL) is shown in **Figure 4**.

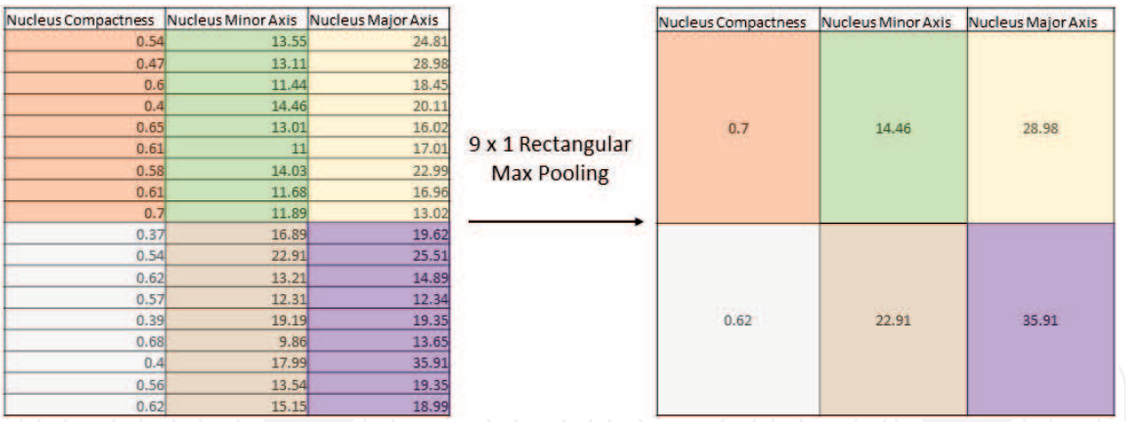


Figure 4.
Max Rectangular Pooling Layer Operation Representation – This shows a sample of how Rectangular Pooling Layer affects our input dataset. With a rectangular pool matrix, we ensure that non-related columns are not pooled together, unlike in a conventional square pooling layer. The transformation is a smaller dataset, with no loss in representation. This reduction in data size, results in faster learning generalisation and computation of the model.

In another experiment comparing a plain vanilla 2D CNN with RPL of 9 x 1 dimension and a conventional square pooling layer (SPL) of 3 x 3 dimension, it showed both having the same max function (**Table 10**). Comparing the training record of both pooling shapes, the RPL generalised at a much faster rate, of about 500 epochs ahead of SPL to achieve the same MAE. The former also achieved a lower MAE at the end of the training. In the context of a large dataset and DL network, using RPL in a non-spatial 2D dataset could achieve significant reduction in computational time.

2.2.2 Validation and evaluation

2.2.2.1 Validation

In the context of medical dataset, one common hampering factor is having a small dataset. This results in a validation process that is not robust enough as there may be an uneven distribution of data across the dataset. Traditional holdout validation is not rigorous enough to negate this effect and may result in an unfair representation of the efficacy of the model. This could be overcome with the use of K-Fold cross validation (K-cv), which is done by splitting the dataset by *k* iteratively holding out the sections of the data and evaluating the model with an average prediction error of all *k* evaluations (**Figure 5**).

Epochs	Rectangular Max Pooling MAE (months)	Square Max Pooling MAE (months)	Difference between Square Max Pooling and Rectangular Max Pooling (% Base using Rectangular Max Pooling)
500	81.341	81.869	0.65
1000	73.678	76.340	3.61
1500	67.018	71.577	6.80
2000	60.403	65.921	9.14
2500	53.717	58.351	8.63
3000	48.375	52.431	8.38
3500	45.761	49.085	7.26

Table 10.
Result of Rectangular Pooling Layer (RPL) vs. Square Pooling Layer (SQL).

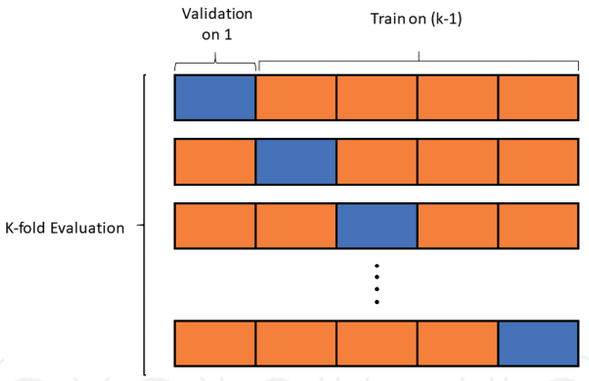


Figure 5.
K - Fold Cross Validation – By splitting our dataset into k folds, we can evaluate our model across the entire dataset independently. This is especially critical for small datasets (as is the case in medical context).

K-cv provides a more robust way of validating a model by validating a model with the entire data set. A study by Rodriguez et al. confirmed that K-cv reduces variance in prediction error and recommends implementing a K-cv whenever computationally possible [39]. In this study, the BC dataset was split into four groups of 23 patients and the standard deviation σ of each group is evaluated. It was discovered that the σ across all four groups was 8.43 months, which is a substantial amount in its ability to misrepresent the efficacy of the dataset.

2.2.2.2 Evaluation

The evaluation step acts as a feedback loop to the development of the CNN model. An iterative approach must be taken to analyse these results from a DL and a medical point of view to understand how further improvements could be made to the CNN model. Firstly, a model was built to evaluate clinical dataset followed by another model to evaluate the cell dataset. In an experiment with 107 patients, an adaptation of Dense ResNet [40] to the clinical data was used. A 2D CNN Wide Residual Network (WRN) [41] was also adapted for the cell data (**Figure 6**). A benchmark was developed on the dataset as a starting point for comparison, as this was a greenfield application. A simple vanilla dense network was used as a starting point to benchmark the results for the clinical dataset containing patients' information such as age, ethnicity, and tumour size. For the immunopathological dataset, we use a benchmark CNN model from the imaging domain as our starting point. MobileNet50 V2 [42] was chosen for the starting benchmark for

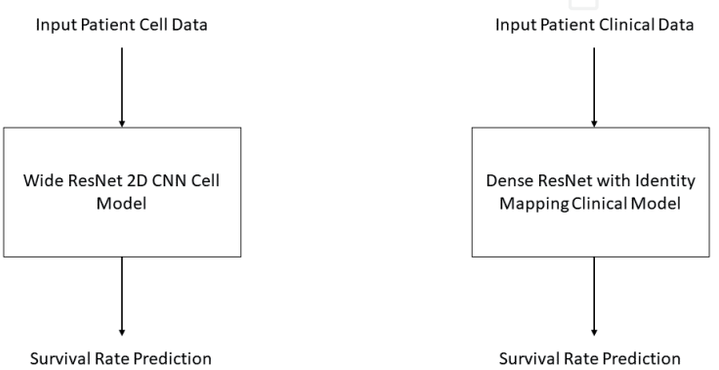


Figure 6.
Proposed Network Layout – Two independent model first learn representation from their respective dataset, which will have their weights combined together to create a unified model to create a single prediction from both datasets.

immunopathological dataset due to its accuracy, and training speed secondary to its small size (**Table 11**).

A sample k-fold training record as shown in **Figure 7**, shows the overfitting tendencies as the training error is minimised, but the validation error is not minimised. This could be attributed to the unsuitability of the models in the imaging domain without adaptation, which emphasises the importance of using the framework to adapt available CNN models to specific needs. In this study, a more suitable model was developed to clean up, augment, and enhance our dataset following the steps of the framework.

As shown in **Table 12**, the clinical dataset results were augmented from ±15.69 to ±8.24 by the immunopathological data from mIHC/IF with two additional information: number of stromal immune cells and cancer cells of each patients quantitated from the cell dataset. The results were subsequently normalised and iteratively developed to form a new Dense Neural Network based on the ResNet architecture that was better suited to the dataset. With regards to the cell dataset, the results

Clinical Vanilla MAE (months)	Cell MobileNet50 V2 MAE (months)	Unified Model MAE (months)
± 15.69	± 81.66	± 97.34

Table 11.
Benchmark Results for Both Dataset.

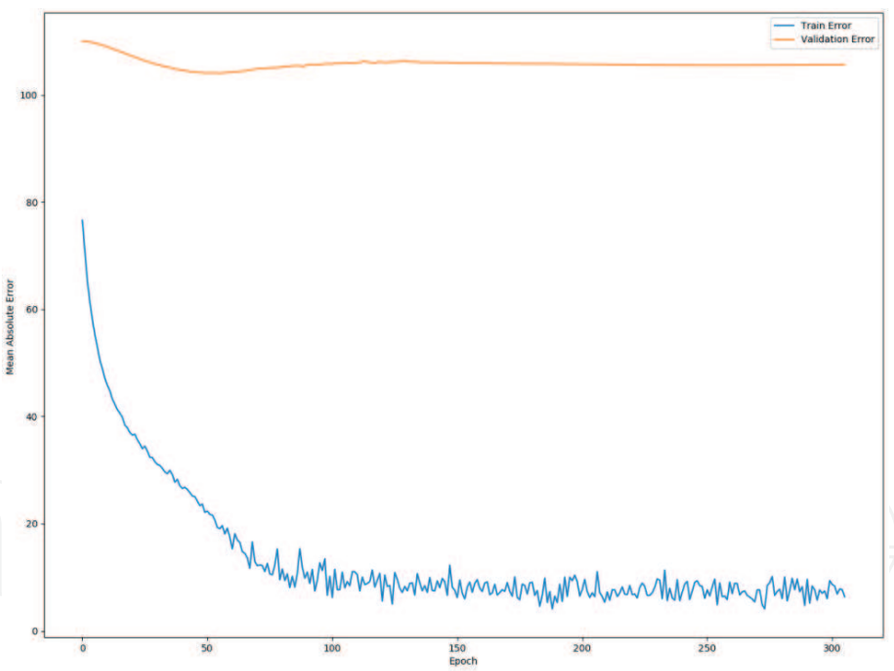


Figure 7.
Training History of MobileNet V2 – Benchmarking using a conventional general CNN model. Without adapting the model from imaging domain to our specific use case, we see a tendency to overfit by the divergence decreasing training error (Blue line) and the constant validation error (red line). This shows that there is no generalisation in the model, which serves as a good starting point, and a reminder of the need to adapt CNN to our specific use case.

Clinical ResNet MAE (months)	Cell WRN MAE (months)	Unified Model MAE (months)
± 8.24	± 40.23	± 55.23

Table 12.
Result for Both Dataset and Unified Model Using Adapted Models which factored in immunopathological data from mIHC/IF.

Survival Rate Cut-Off (Months)	No. of Patients	Combined MAE (Months)	Clinical + immunopathological data
Full Dataset (Benchmark)	107	± 55.23	± 8.24
≥12	96	± 52.17	± 8.78
≥16	92	± 35.11	± 10.67
≥20	88	± 25.86	± 11.42

Table 13.
MAE of dataset of Cut-Off Survival Rate.

were normalised to develop a more suitable CNN using WRN with RPL. Significant improvements in both the cell and clinical dataset were seen, which is appended with immunopathological data. The results were further improved by including a threshold based on the patients' survival rate.

A filter of patients with lower survival rate was experimented where the dataset was split with an arbitrary cut-off of ≥12 months, ≥16 months, and ≥ 20 months. Evaluating using the same unified model, **Table 13** showed the following MAE on 5-fold K-cv for each cut-off, where comparison of the combined clinical dataset and cell dataset were made after applying cut-off filter. The clinical dataset augmented with the stroma and tumour count from the cell dataset is also reflected in **Table 13** for reference. The increased in cut-off threshold meant that the model had a smaller dataset. Therefore, an increased in MAE of the model was expected, which was in line with the results shown in **Table 13**.

3. Limitations

Some limitations of this study should be noted. Firstly, this study uses a small dataset, which meant that the results could be less robust and of a lower confidence level. Although, this was minimised with the use of k-fold cross validation, more advanced techniques such as semi-supervised learning could be explored to augment the dataset. Secondly, there is currently no medical evidence to support using a cut-off to segregate patients as a valid approach. The approach used in this study is solely from a DL standpoint and therefore requires more medical based research to prove its validity. Moreover, given the novelty of the proposed framework, there is currently limited literature to support its application in other medical domains.

4. Conclusions

The adaptation of DL technology with the use of mIHC in the analysis of complex data is in the upcoming alternative approach of analysis in the field of immunopatholgy. However, given its novelty, further studies are needed to optimised the framework to enable application in varies medical field. Nevertheless, the framework proposed in this chapter serves to provide a starting foundation for application in clinical studies.

Author contributions

Conceptualization and design, Y. Chua and J. Yeong; literature review, S. Goh and Y. Chua; writing-original draft, S. Goh and Y. Chua; intellectual input and

critical review, J. Lee, J. Yeong and Y. Cai.; writing-review and final editing, S. Goh and J. Lee. All authors have read and agreed to the published version of the manuscript.

IntechOpen

Author details

Siting Goh[1], Yueda Chua[4], Justina Lee[3], Joe Yeong[1,2,3*] and Yiyu Cai[4]

1 Division of Pathology, Singapore General Hospital, Singapore

2 Duke-NUS Medical School, Singapore

3 Institute of Molecular Cell Biology (IMCB), Agency of Science, Technology and Research (A*STAR), Singapore

4 Nanyang Technological University (NTU), Singapore

*Address all correspondence to: yeongps@imcb.a-star.edu.sg

IntechOpen

© 2021 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. CC BY

References

[1] Xin Yao, "Evolving artificial neural networks," in *Proceedings of the IEEE*, vol. 87, no. 9, pp. 1423-1447, Sept. 1999, doi: 10.1109/5.784219.

[2] Deng, L., & Yu, D. (2014). Deep learning: methods and applications. Foundations and trends in signal processing, *7*(3-4), 197-387.

[3] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, *521*(7553), 436-444.

[4] Bera, K., Schalper, K. A., Rimm, D. L., Velcheti, V., & Madabhushi, A. (2019). Artificial intelligence in digital pathology—new tools for diagnosis and precision oncology. Nature reviews Clinical oncology, *16*(11), 703-715.

[5] Elmore, J. G., Longton, G. M., Carney, P. A., Geller, B. M., Onega, T., Tosteson, A. N., ... & O'Malley, F. P. (2015). Diagnostic concordance among pathologists interpreting breast biopsy specimens. Jama, *313*(11), 1122-1132.

[6] Bejnordi, B. E., Veta, M., Van Diest, P. J., Van Ginneken, B., Karssemeijer, N., Litjens, G., ... & Geessink, O. (2017). Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. Jama, *318*(22), 2199-2210.

[7] Chen, J., & Srinivas, C. (2016). Automatic lymphocyte detection in H&E images with deep neural networks. *arXiv preprint arXiv:1612.03217*.

[8] Cruz-Roa, A., Gilmore, H., Basavanhally, A., Feldman, M., Ganesan, S., Shih, N., Tomaszewski, J., Madabhushi, A., & González, F. (2018). High-throughput adaptive sampling for whole-slide histopathology image analysis (HASHI) via convolutional neural networks: Application to invasive breast cancer detection. PloS one, *13*(5), e0196828. https://doi.org/10.1371/journal.pone.0196828

[9] Tan, W., Nerurkar, S. N., Cai, H. Y., Ng, H., Wu, D., Wee, Y., Lim, J., Yeong, J., & Lim, T. (2020). Overview of multiplex immunohistochemistry/immunofluorescence techniques in the era of cancer immunotherapy. Cancer communications (London, England), *40*(4), 135-153. https://doi.org/10.1002/cac2.12023

[10] Gong J, Chehrazi-Raffle A, Reddi S, Salgia R. Development of PD-1 and PD-L1 inhibitors as a form of cancer immunotherapy: a comprehensive review of registration trials and future considerations. J Immunother Cancer. 2018 Jan 23;6(1):8.

[11] Hellmann MD, Ciuleanu T-E, Pluzanski A, Lee JS, Otterson GA, Audigier-Valette C, et al. Nivolumab plus Ipilimumab in Lung Cancer with a High Tumor Mutational Burden. New England Journal of Medicine. 2018;378(22):2093-2104.

[12] Yau T, Zagonel V, Santoro A, Acosta-Rivera M, Choo SP, Matilla A, et al. Nivolumab (NIVO) + ipilimumab (IPI) + cabozantinib (CABO) combination therapy in patients (pts) with advanced hepatocellular carcinoma (aHCC): Results from CheckMate 040. Journal of Clinical Oncology. 2020;38(4_suppl):478-478.

[13] Garon EB, Rizvi NA, Hui R, Leighl N, Balmanoukian AS, Eder JP, et al. Pembrolizumab for the Treatment of Non–Small-Cell Lung Cancer. New England Journal of Medicine. 2015;372(21):2018-2028.

[14] Eggermont AMM, Blank CU, Mandala M, Long GV, Atkinson V, Dalle S, et al. Adjuvant Pembrolizumab versus Placebo in Resected Stage III Melanoma. N Engl J Med. 2018 May 10;378(19):1789-1801.

[15] Coons AH, Creech HJ, Jones RN. Immunological Properties

of an Antibody Containing a Fluorescent Group. Proceedings of the Society for Experimental Biology and Medicine. 1941 1941/06/01 [cited 2020/07/13];47(2):200-202.

[16] Ahmadzadeh M, Johnson LA, Heemskerk B, Wunderlich JR, Dudley ME, White DE, et al. Tumor antigen-specific CD8 T cells infiltrating the tumor express high levels of PD-1 and are functionally impaired. Blood. 2009 Aug 20;114(8):1537-1544.

[17] Muenst S, Hoeller S, Willi N, Dirnhofera S, Tzankov A. Diagnostic and prognostic utility of PD-1 in B cell lymphomas. Dis Markers. 2010;29(1): 47-53.

[18] Yeong J, Lim JCT, Lee B, Li H, Chia N, Ong CCH, et al. High Densities of Tumor-Associated Plasma Cells Predict Improved Prognosis in Triple Negative Breast Cancer. Front Immunol. 2018;9:1209.

[19] Hainaut P, Plymoth A. Targeting the hallmarks of cancer: towards a rational approach to next-generation cancer therapy. Curr Opin Oncol. 2013 Jan;25(1):50-51.

[20] Kim S-W, Roh J, Park C-S. Immunohistochemistry for Pathologists: Protocols, Pitfalls, and Tips. Journal of Pathology and Translational Medicine. 2016 10/13 04/18/received 08/05/rev-recd 08/08/accepted;50(6):411-418. Available from: ScienceCentral

[21] Curigliano G, Burstein HJ, Winer EP, Gnant M, Dubsky P, Loibl S, et al. De-escalating and escalating treatments for early-stage breast cancer: the St. Gallen International Expert Consensus Conference on the Primary Therapy of Early Breast Cancer 2017. Ann Oncol. 2017 Aug 1;28(8):1700-1712.

[22] Tay TKY, Thike AA, Pathmanathan N, Jara-Lazaro AR, Iqbal J, Sng ASH, et al. Using computer assisted image analysis to determine the optimal Ki67 threshold for predicting outcome of invasive breast cancer. Oncotarget. 2018;9(14):11619-11630. Available from: PubMed

[23] Koopman T, Buikema HJ, Hollema H, de Bock GH, van der Vegt B. Digital image analysis of Ki67 proliferation index in breast cancer using virtual dual staining on whole tissue sections: clinical validation and inter-platform agreement. Breast Cancer Res Treat. 2018 May;169(1):33-42.

[24] Yeong J, Tan T, Chow ZL, Cheng Q, Lee B, Seet A, et al. Multiplex immunohistochemistry/immuno-fluorescence (mIHC/IF) for PD-L1 testing in triple-negative breast cancer: a translational assay compared with conventional IHC. J Clin Pathol. 2020 Jan 22.

[25] Halse H, Colebatch AJ, Petrone P, Henderson MA, Mills JK, Snow H, et al. Multiplex immunohistochemistry accurately defines the immune context of metastatic melanoma. Scientific Reports. 2018 2018/07/24;8(1):11158.

[26] Lu S, Stein JE, Rimm DL, Wang DW, Bell JM, Johnson DB, et al. Comparison of Biomarker Modalities for Predicting Response to PD-1/PD-L1 Checkpoint Blockade: A Systematic Review and Meta-analysis. JAMA Oncol. 2019 Jul 18;5(8):1195-1204.

[27] Johnson DB, Bordeaux J, Kim JY, Vaupel C, Rimm DL, Ho TH, et al. Quantitative Spatial Profiling of PD-1/PD-L1 Interaction and HLA-DR/IDO-1 Predicts Improved Outcomes of Anti-PD-1 Therapies in Metastatic Melanoma. Clin Cancer Res. 2018 Nov 1;24(21):5250-5260.

[28] Giraldo NA, Nguyen P, Engle EL, Kaunitz GJ, Cottrell TR, Berry S, et al. Multidimensional, quantitative assessment of PD-1/PD-L1 expression in patients with Merkel cell carcinoma

and association with response to pembrolizumab. J Immunother Cancer. 2018 Oct 1;6(1):99.

[29] Fiore C, Bailey D, Conlon N, Wu X, Martin N, Fiorentino M, et al. Utility of multispectral imaging in automated quantitative scoring of immunohistochemistry. Journal of Clinical Pathology. 2012;65(6):496-502.

[30] Abel EJ, Bauman TM, Weiker M, Shi F, Downs TM, Jarrard DF, et al. Analysis and validation of tissue biomarkers for renal cell carcinoma using automated high-throughput evaluation of protein expression. Hum Pathol. 2014;45(5):1092-1099.

[31] Feng Z, Bethmann D, Kappler M, Ballesteros-Merino C, Eckert A, Bell RB, et al. Multiparametric immune profiling in HPV– oral squamous cell cancer. JCI Insight. 2017 07/20/;2(14).

[32] Mascaux C, Angelova M, Vasaturo A, Beane J, Hijazi K, Anthoine G, et al. Immune evasion before tumour invasion in early lung squamous carcinogenesis. Nature. 2019 Jun 26.

[33] Parra ER, Francisco-Cruz A, Wistuba, II. State-of-the-Art of Profiling Immune Contexture in the Era of Multiplexed Staining and Digital Analysis to Study Paraffin Tumor Tissues. Cancers (Basel). 2019 Feb 20;11(2).

[34] Tan AS, Yeong JPS, Lai CPT, Ong CHC, Lee B, Lim JCT, et al. The role of Ki-67 in Asian triple negative breast cancers: a novel combinatory panel approach. Virchows Arch. 2019 Dec;475(6):709-725.

[35] Yeong J, Lim JCT, Lee B, Li H, Ong CCH, Thike AA, et al. Prognostic value of CD8 + PD-1+ immune infiltrates and PDCD1 gene expression in triple negative breast cancer. Journal for ImmunoTherapy of Cancer. 2019 2019/02/06;7(1):34.

[36] Tan WCC, Nerurkar SN, Cai HY, Ng HHM, Wu D, Wee YTF, et al. Overview of multiplex immunohistochemistry/immunofluorescence techniques in the era of cancer immunotherapy. Cancer Commun (Lond). 2020 Apr;40(4):135-153.

[37] Yeong J, Thike AA, Lim JC, Lee B, Li H, Wong SC, Hue SS, Tan PH, Iqbal J. Higher densities of Foxp3+ regulatory T cells are associated with better prognosis in triple-negative breast cancer. Breast Cancer Res Treat. 2017 May;163(1):21-35.

[38] Z. Jin Yan, "INVESTIGATION OF ARTIFICIAL INTELLIGENCE FOR MEDICAL IMAGE BASED DIAGNOSIS USING DEEP LEARNING," B.Eng Aerospace Engineering Final Year Project, Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore 2018.

[39] J. D. Rodriguez, A. Perez, and J. A. Lozano, "Sensitivity analysis of k-fold cross validation in prediction error estimation," IEEE transactions on pattern analysis and machine intelligence, vol. 32, no. 3, pp. 569-575, 2009.

[40] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778.

[41] S. Zagoruyko and N. Komodakis, "Wide residual networks," *arXiv preprint arXiv:1605.07146,* 2016.

[42] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4510-4520.