Università degli Studi di Napoli

"Federico II"

SCUOLA DI MEDICINA E CHIRURGIA

Dipartimento di Scienze Biomediche Avanzate

Dottorato di Ricerca in Scienze Biomorfologiche e Chirurgiche

XXXIV ciclo

A deep learning model to predict Ki-67 positivity in Oral Squamous Cell Carcinoma

Tutor:

Ch.mo. Prof.

Stefania Staibano

Coordinatore del Dottorato:

Ch.mo. Prof.

Alberto Cuocolo

Candidato:

Francesco Martino

ANNO ACCADEMICO 2020/2021

Contents

1 Introduction **3**

1.1 AI: The ”Third Revolution” in Anatomic Pathology . . . 3

1.2 Oral Squamous Cell Carcinoma . . . 4

1.2.1 Epidemiology and risk factors . . . 4

1.2.2 Staging, clinical presentation and therapy . . . 4

1.3 Digital Pathology . . . 5

1.4 Artificial intelligence . . . 5

1.4.1 Computer Vision . . . 6

1.5 Machine and Deep Learning . . . 8

1.5.1 Classification and Regression . . . 9

1.5.2 Neural Networks . . . 10

1.5.3 Metrics . . . 12

1.5.4 Loss and Gradients . . . 13

1.5.5 Datasets . . . 14

1.5.6 Generative networks . . . 14

1.6 Machine/Deep Learning in Digital Pathology . . . 16

1.6.1 AI applications in Oral Squamous Cell Carcinoma analysis . . . 17

2 Aim of the study **18**

3 Materials and methods **19**

3.1 Study Population . . . 19

3.2 Virtual Staining . . . 19

3.2.1 Dataset generation . . . 19

3.2.2 Pix2Pix implementation . . . 20

3.2.3 IHC Quantification . . . 20

3.2.4 Android application . . . 21

3.2.5 Code and Statistical Analysis . . . 21

4 Results **22**

4.1 Dataset Generation . . . 22

4.2 Synthetic Images Likelihood . . . 24

4.3 Synthetic and real IHC concordance . . . 27

5 Discussion and conclusion **30**

Abstract

Anatomical Pathology is living its third revolution, facing a radical transformation from analogical pathology to digital pathology, with the new artificial intelligence applications becoming part of the clinical practice. Other than classification, detection, and segmentation models, the spotlight is on predictive models which may impact not only on diagnostic procedures but also on laboratory activity, reducing the usage of consumables and the turn-around time. In our study, we aimed to develop a deep learning model capable of generating a synthetic Ki-67 immunohistochemestry using H&E images as input. To develop our model, we retrieved 175 Oral Squamous Cell Carcinoma from the archives of the Pathology Unit of the University Federico II, and built four TMAs. We then generated one slide from each TMA which was in first instance stained with H&E protocol and subsequently destained and restained using anti-Ki-67 immunohistochemistry. Cores were then dearreyed and tiled to create a dataset to train a Pix2Pix model to convert H&E images to IHC. Our model resulted in realistic synthetic images, as pathologists were able to recognise the syntetic images only in half of the cases. Then, we quantified IHC positivity using QuPath, achieving high levels of concordance between real IHC and synthetic IHC. Moreover, a categorical analysis using three cutoff of Ki-67 positivity (5%, 10%, and 15%) showed high positive predicted values. Overall, although these results need be confirmed on a larger dataset in a multicentric environment, our model represents a promising tool to gather Ki-67 positivity information directly on H&E slides, reducing the laboratory demand and improving the management of patients, being also a valid opportunity for smaller hospitals which cannot keep up with the raising necessity of several immunohistochemical staining.

1 Introduction

1.1 AI: The "Third Revolution" in Anatomic Pathology

Nowadays, pathology is quite different from its dawn, since it underwent through two major revolutions and it is living its third revolution, the AI era [1]. Indeed, pathology's origin goes back the first half of 19^{th} century with Karl von Rokitansky and Rudolf Virchow, as the first is supposed to have performed over 30,000 autopsies, while the latter introduced the microscopical analysis of tissues, leading to the cell theory of disease [2]. At that time, and for a long time, anatomical pathology was about the morphological evaluation of images to assess the diagnosis, without any information about proteins or gene expression, until in 1970 Sternberger et al. described the application of immunohistochemistry to tissues, using the peroxidase/antiperoxidase system in a paper cited more than 6000 times, revolutionising the anatomical pathology [2], [3].

Then, in 70s, immunohistochemistry rapidly became part of anatomical pathology workflow conducing to more accurate diagnosis and enabling the role of pathologists in prognosis prediction thanks to the analysis of proteins expression, along with the diagnosis itself. Furthermore, in 2001, another milestone has been achieved in anatomical pathology: Imatinib, a small inhibitor molecule, was granted the FDA approval for the treatment of patient affected by Chronic Myelogenous Leukemia positive to Philadelphia chromosome [4], paving the way to the second revolution in anatomical pathology which, through the years, reshaped from morphological pathology to morphomolecular pathology, with an increasing interest in molecular analysis guided by the invention of highly efficient techniques as the Next Generation Sequencing [5]. In that framework, pathologist have been asked to outdo the diagnosis itself and to perform a comprehensive analysis of patients, enriching the diagnosis with a molecular characterisation addressing the prognosis and the therapy too.

However, the history of pathology is defined, as anticipated, by another revolution that changed again the workflow and the activity in anatomical pathology: digital pathology and whole slide images (WSI), a glass slide through digitised a dedicated scanner.

Paradoxically, digital pathology is far older than expected and the first "Virtual Microscope" goes back to 1997 and it was a software intended to emulate the behaviour of a physical microscope [6], while in 1994 James Bacus designed the first Whole Slide Scanner and his two patents were filed in 1997 and 1998 [7]. During following years, several scanners were introduced and multiple roles for WSI become well established (i.e. telepathology [8] or education [9]), but their integration in routine workflow was still far away. Then, in 2011, Isaacs et al. presented a valuable example of integration of WSI in clinical practise [10], setting a new stage in anatomical pathology, with many hospital rapidly taking a role as actor in the new panorama of digitised surgical pathology [11]–[14]. Eventually, thanks to the rapid growth of digitised wards, the

technological improvement in slides scanners, and the advances in the field of computer science, anatomical pathology has been hauled by the introduction of artificial intelligence to automatically analyse WSI, with many application to most common cancers (e.g. breast, prostate) and, recently, even to less common ones, as the Oral Squamous Cell Carcinoma.

1.2 Oral Squamous Cell Carcinoma

1.2.1 Epidemiology and risk factors

Squamous cell carcinoma (SCC) represents 95% of Head and Neck Cancers and the sixth most common malignancy in the world [15]. Head and Neck Squamous Cell Carcinoma (HNSCC) mainly occurs in male patients, with a ratio of 3:1 against women, between the fifth and the sixth decades of live, with a country-related variability [16], while oropharyngeal (OP) cancers which may occur during the forth decade of life [17]. Moreover, the incidence of OP cancer appears to be increasing due to the raise of HPV-related cancer, which nowadays accounts for approximately 70% of all oropharyngeal cancers in the USA [17]. HNSCC results in nearly 640,000 cases/year worldwide among men and 210,000 cases/year for female, with an estimated incidence of 16.1 cases per 100,000 people per year [18]. Although great advances have been achieved in surgery and radiotherapy, HNSCC mortality still remains close to 50% [19], and HNSCC insurgence is highly influenced by environmental factors as tobacco, both smoked and chewed, alcohol consumption, and viral infection by High-Risk Human Papillomavirus (HPV) genotypes [20]. Oropharyngeal squamous cell carcinoma are indeed subdivided in HPV-Positive and HPV-Negative cancers, and their outcome is different among the two groups, with the HPV-Positive cancers showing a better prognosis compared to negative ones [21].

1.2.2 Staging, clinical presentation and therapy

Staging is among the most determinant factors for the prognosis and the treatment choice of patients diagnosed with HNSCC. tumour staging is performed according to TNM (tumour-Node-Metastasis), considering the size of the tumour, the lymph nodes involvement, and the presence of metastases, respectively, according to the last edition of the American Joint Committee on Cancer (AJCC) guidelines [22].

From a clinical perspective, leukoplakia and erythroplakia represents precancerous lesions which may progress into invasive cancers [23] and the risk of progression depends on the size of the lesion and, most importantly, the presence of dysplasia [24]. HNSCC treatment depends on the stage and the site of occurrence. Stage I and II tumours are treated with surgery and/or radiotherapy, whole advanced stages (III and IV) need a more aggressive approach, combining surgery, radiation and chemotherapy [25]. Up to date, we still lack of reliable prognostic and therapy-response predictive biomarkers, along with effective therapeutic targets, hence much effort is still needed to fully understand the biology of these tumours in order to improve the

clinical management of HNSCC patients. To pursue this goal, research in pathological anatomy makes use of the entire classical repertoire of tools for studying solid tumours, to which Digital Pathology has recently been added, with the promise to provide a valuable support.

1.3 Digital Pathology

Digital Pathology is defined as *"A dynamic, image-based environment that enables the acquisition, management and interpretation of pathology information generated from a digitised glass slide. Often used interchangeably with "Virtual Microscopy.""* by the Digital Pathology Association [26], and relies on Whole Slide Images (WSI) acquired using a slide scanner. Whole Slide Images are extremely-high resolution images, with a typical size if 100,000×100,000 of pixels and a weight in the order of gigabytes [27], obtained by high-speed scanners which exist in a wide range of assets, ranging from single slide scanners to 400 and more slides scanners, or from linear scanning to CCD-camera based scanners. The introduction of WSI in surgical pathology started a revolution that affected not only slide analysis (i.e. immunohistochemistry (IHC) signal quantification or cell count) but also routine diagnostic procedures [28]. Indeed, several hospitals around the world converted to a fully-digital workflow [14] after that various studies reported the validity of virtual slides as diagnostic tool, confirming the agreement between pathologists using digital slides and physical slides, and demonstrating the non-inferiority of WSI compared with physical slides [29]. Moreover, WSI are relatively easily stored and are easily remotely consulted via a shareable link which facilitates the collaboration between pathologist settled in different locations [30]. Furthermore, due to their digital nature and the huge amount of pixels, WSI enabled computer vision approaches to pathological images, allowing to perform classical computer vision approaches, as cell detection or image classification and segmentation, and advanced ones, as molecular prediction and 3D reconstruction, applying Artificial Intelligence to histopathological slides [31].

1.4 Artificial intelligence

Artificial intelligence and *"robotics"* are, in spite of common sense, among the oldest ideas in humankind and have their roots in ancient mythology as the Golem, a Judaic traditional being made of inanimate matter, and many other examples exist all along history [32]. The attempt to create a machine capable of human thoughts has even lead to crafty methods to exhibit intelligent machines, as *the Turk*, a 18th centuries artificial-intelligent chess player that turned out to be a fake automaton secretly manoeuvred by well trained human player [33]. However, the first actual instance of artificial intelligence is *Enigma*, the Turing's machine [34], that was subsequently followed by the ideation of the concept of Neural Network (NN) by Walter Pitts and Warren McCulloch [35]. Then, one of Pitts and McCulloch students, Marvin Minsky, together with Dean Edmonds, built the first NN, namely Stochastic neural analog reinforcement

calculator (SNARC) [36].

Years later, John McCarthy organised the event proposed by Minsky, with Rochester and Shannon, *Dartmouth Summer Research Project on Artificial Intelligence*, which is considered the proper birth of the Artificial Intelligence [37]. After this event, AI rapidly developed and several tasks were approached, with a particular interest in board and video-games due to their scoring system (as count of won matches or earned points) which facilitate the evaluation of the AI system learning improvement [38]. Other than games, AI has been used to accomplish classification, regression, optimisation, and other application as text translation or generation, with both supervised and unsupervised approaches. [39]

1.4.1 Computer Vision

Following the development of AI, computational approaches has been used to handle with images, resulting in the birth of the so-called "Computer Vision". Computer vision (CV) is the attempt to retrieve information from images using computers, or, in other words, to automatically analyse images as an human would do, but using a computer. Computer Vision goes back to late fifties, when Hubel and Wiesel published a paper describing the discovery that simple cat's cortical neurons get activated by sharp lines and edges, which is the core principle in CV and Deep Learning (DL) [40]. Then, in the same years, Russell Kirsch et al. developed the first image scanner and their first scan, Russell Kirsch's son, Walden Russel, is now exposed in the Portland Art Museum. They also applied some basic edge-detection algorithms and later in seventies Kirsch devised the "Kirsch operator", one of the most famous edge detection algorithms [41]. The academic interest to CV raised when Larry Roberts presented the possibility to derive 3D information from 2D perspective views of objects [42]. Hence, the attention on CV started getting higher and higher leading to more sophisticated approaches to image analysis as Machine Learning (ML) and Deep Learning (DL). ML and DL relies on the idea of letting computer learn information from a dataset in order to perform tasks as classification and segmentation on new images. Moreover, due to the nature of images, a matrix of bytes, every kind of image (i.e. photos, microscopic or telescopic images, X-Rays, CT, virtual histological slides, etc.) can be used to develop new algorithms for different aims, attracting the interest from several field, both medical and non medical. Hence, a brief glossary of terms is shown in Table 1.

Term	**Short**	**Meaning**
Artificial Intelligence	AI	The idea that computer can act as human being would in a certain situation, because programmed or trained to.

Term	Short	Meaning
Computer Vision	CV	The application of algorithm to allow the computer to "watch" images, perform specific task (i.e. classification, detection, segmentation), and predict an output.
Prediction		The output of an algorithm. Algorithms use data to perform an estimation of the output and, then, a prediction.
Machine Learning	ML	A sub-field of artificial intelligence, consisting in a series of models which allow the computer to learn how to accomplish a task from data, without direct instruction about a specific task.
Deep Learning	DL	A sub-field of machine learning based on the application of artificial neural networks to learn from data and perform tasks.
Metrics		Measures of the effectiveness of a model, used to evaluate its reliability.
Loss		The measure of the effectiveness of the training procedure, and it is calculated as the difference between the predicted and the actual value of an element through a Loss Function
Datasets		In order to learn information from data, datasets must be provided to machine learning models. A training datasets is used to provide information to the model, a test set is to assess the model quality during training, while an external validation set is used to verify to generalizability of the model
Training		The procedure which allows the model to calculate its weights to predict an output given input data
Testing		The procedure of testing the model during the training to recalculate its weights and to improve the quality of the model
Validation		The procedure of validation of the model using never-seen data, to asses how the model accomplish with new samples and to verify if the model is under- or overfitted to training and test datasets

Term	Short	Meaning
Underfitting		The situation in which validation performance are higher than testing ones. The model is learning less than it could, hence the model is too simple.
Overfitting		The situation in which validation performance are lower than testing ones. The model is over fitted to training data and is not generalizable. It may depend by a over-complex model or a unbalanced dataset
Classification		The task of sorting elements in classes
Object Detection		The task of identify the bounding box of elements in an image
Segmentation		The task of identify the precise shape of an element of interest in a image
Regression		The task of predicting a continuous value given an input
Artificial Neural Networks	ANN	Algorithms made of artificial neurons, capable of learning from data and perform artificial intelligence tasks.
Artificial Neurons		They are the fundamental unity of ANN. Each neuron has an input which is processed by an Activation Function resulting in an output which can be sent to other neurons.
Activating Function		A mathematical function used to determine the output of a neuron given an input
Generative Adversarial Networks	GAN	A particular kind of network, made by two adversarial network (a generator and a discriminator), competing each other to generate images and to discriminate real from synthetic images.

Table 1: A glossary of Artificial Intelligence and Deep Learning terms

1.5 Machine and Deep Learning

Machine and deep learning are the processes allowing machine to learn, with a multi-step and iterative process. Generally, letting a machine learns means minimising the *loss* improving its *metrics* through the optimisation of its hyper-parameters in order to calculate and define the correct weight of the algorithm, learning how to perform classification, segmentation, or regression tasks.

1.5.1 Classification and Regression

Classification is the attempt to subdivide data in different categories (namely *labels*), while regression is meant to predict the value of an element according to its features, and they may be achieved using both supervised (training with given labels) and unsupervised (without given labels) algorithms [43]. Supervised algorithms in data classification and regression include K-Nearest Neighbours, Decision Trees, Random Forests, Support Vector Machines (SVM), Naive Bayes, and Neural Networks [43].

- **K-Nearest Neighbours** is a supervised learning algorithm intended to perform class prediction or regression according to similarity of a data-point to ground truth data references. It is based on the idea that same class objects share common features and then it is possible to define a data-point's class identifying the most similar object, with the similarity defined as the Euclidean distance between the data-point's features and the references one. Each data-point will thus be assigned to the same class of the nearest object. Other then distance, the algorithms accounts also for the K parameter that represents the K number of nearest object to compare in order to classify the data-point as the most represented class among the K elements. [44]

- **Decision Tree** are non-parametric supervised learning methods. Their aim is to predict the value of a target variable by decision rules from its features. Decision tree algorithm recursively iterate and split through all the data-point features calculating the *accuracy cost* and keeping the split with the lowest cost. Overall, the model is optimised until all the splits which cost the least are kept. [45]

- **Random Forest** algorithms are an ensemble of multiple decision trees and they output the modal predicted class/value by the back-end decision trees, overcoming their attitude to overfitting [46]

- **Support Vector Machines** are supervised classification and regression algorithms intended to identify an *hyperplane* in N-dimensional space to classify given data-points. The best separation is achieved with the hyperplane(s) with the largest distance from the nearest elements of each label. These margin elements are the so called *support vectors* which the machine is named after. [47]

- **Naive Bayes** is a classification algorithm based on the probabilistic assignment of a label namely the probability $p(A|B)$ of the event A given the event B, defined as the ratio between the probability $p(B|A)$ of B given A, accounting for the probabilities of the event A $p(A)$ and B $p(B)$ alone, so that:

$$p(A|B) = \frac{p(B|A)p(A)}{p(B)} \tag{1}$$

with A being the classification label and B the features. [48]

- **Neural Networks** are algorithms that mimic the human neurons behaviour to accomplish different tasks as classification, regression along with other more complex applications and represent the basis for deep learning.

Unsupervised learning is a group of algorithms intended to perform data classification without the need of labelled data, and although less accurate and trustworthy, they are frequently used because of the difficulty to collect labelled data. Most common unsupervised learning algorithms are Principal Component Analysis (PCA) and Clustering ones.

- **Principal Component Analysis** is a statistical method to convert a series of observations into *principal components*, namely values of uncorrelated variables.

- **Clustering** is indeed a class of procedures to group objects on the basis of their similarity. There are several types of clustering tools, i.e. hierarchical, which clusters data-points grouping closer ones together, or k-means clustering, where objects are classified according to their closeness to a centroid [49]

All these approaches have been proposed and used for image classification too, but images may be considered as themselves a repository of complex information, hence the branch of AI devoted to image analysis is considered as a nearly separated field called *Computer Vision*, often using Neural Networks.

1.5.2 Neural Networks

As mentioned beforehand, NN have been conceptualised by Pitts and McCulloch and firstly realised by Minsky and Edmonds in 1951. Neural Networks resemble the structure of human neurons having dendrites (**inputs**), a body (**bias**), an axon (**activating function**), and axon terminals (**outputs**), repeated for a number N for each layer, which, in turn, are arranged in a multi-layer organisation composed of an *Input* layer, one or multiple *Hidden* layer(s), and an *Output* layer [Figure 1].

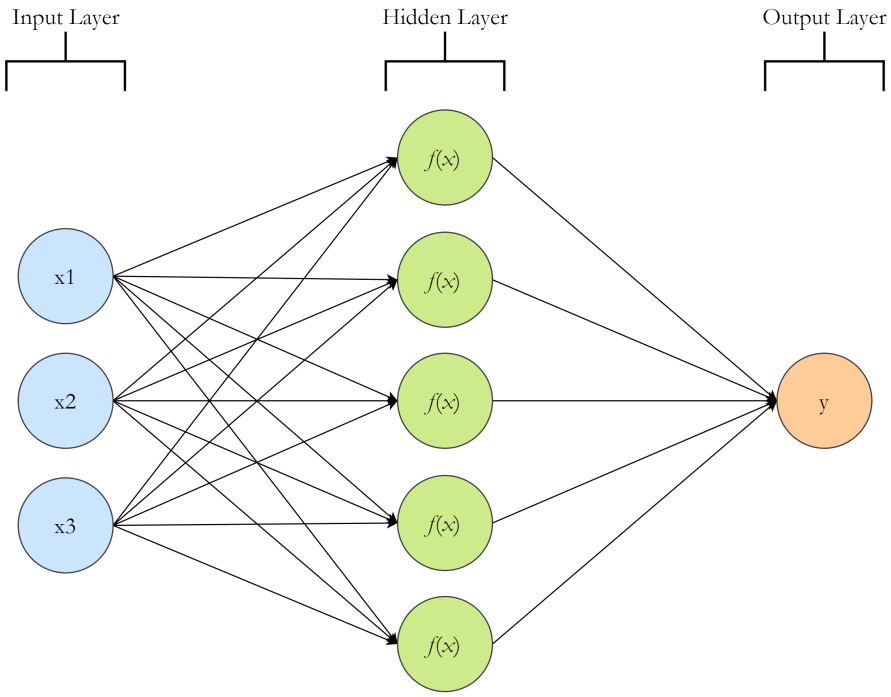

Figure 1: A schematic representation of a Neural Networks. The first layer is named the Input layer and has one neuron for each feature. Inputs are then transmitted to Hidden Neurons in the Hidden Layer, where a mathematical function computes an output and each neuron send its output to the Output Neuron. The output neuron weights and combines all its inputs in a final output, the prediction.

Each neuron represents a mathematical function that takes the inputs and a weight, which is usually initialised with a value of 1 and is recalculated for each training epoch, and then translates these value into a single output that is then used as an input for the subsequent layer of neurons. There are three main layer classes with different roles and each one is necessary to the functioning of NN. Input layer is defined as a *passive layer* as it does not modify data but just copies the inputs and shunts them to each neuron of the first Hidden layer. Here, inputs are processed and each neuron transmits its outputs to the next layer, repeating this process for every hidden layer present in the network architecture. Last hidden neurons will then send their output to the Output layer, where the N neuron will further process data to obtain the N final outputs corresponding the N probabilities to belong to each of the N classes present in the training dataset [50]. Overall, the training process mainly depends on the decision of the proper architecture and the most appropriate loss and metric functions, other than an adequate dataset.

1.5.3 Metrics

Different and variegate metrics exist to evaluate machine learning methods, and it is important to apply the correct one to guarantee a reliability in a real world application. **Accuracy** is the most known classification metrics and is defined as the percentage of correct classification over the total number of classification, but it can be used only for balanced distribution since in case of highly unbalanced distributed data it could result with a value which nearly matches the most represented class frequency, not resembling the effectiveness. To overcome this limit, **precision** is a class specific metrics allowing to calculate the number of correctly predicted samples over the total number of samples belonging to the same actual class. Another important metric is the **recall**, a measurement of correctly predicted labels among all the same predicted class. Precision and recall may be combined in a single score, called **F1 Score**, calculated as:

$$\frac{2 \cdot Precision \cdot Recall}{(Precision + Recall)} \tag{2}$$

Indeed, picking the correct metric during model development and validation, is fundamental to guarantee reliability in a real world data, seeing that a tool to foresee the classification label of a low prevalence element would achieve a close-to-hundred percent accuracy and precision even missing almost every positive element. To better elucidate the differences between metrics and provide a clear overview, a fictitious case study is showed in Table 2.

		Element		
		Present	Absent	Total
Test	Positive	9	1	10
	Negative	90	9900	9990
	Total	99	9991	10000
Metric	Formula			Value
Accuracy	$\frac{TP+TN}{TP+FP+TN+FN}$			0.99
Recall (Sensitivity)	$\frac{TP}{TP+FN}$			0.09
Specificity	$\frac{TN}{FP+TN}$			0.99
Precision (PPV)	$\frac{TP}{TP+FP}$			0.90
F1 Score	$\frac{2 \cdot Precision \cdot Recall}{(Precision+Recall)}$			0.165

Table 2: A fictitious case study to investigate how different metrics describe a classifier effectiveness. In a similar situation (i.e. a test to identify patients positive to an infectious disease with a prevalence of 1%), a test with an accuracy of 0.99 would be totally unreliable as only 9% of positive patients have been identified (and 90 out of 99 positive ones would act as spreader of a disease). Hence, recall and F1 score would have better described the (un)reliability of the test.

The typical output of a model, however, is not a label but indeed a value between 0 and

1 representing the probability of being of a specific class, hence it is necessary to identify a threshold for positive/negative definition. **Receiver Operating Characteristics** curve analysis shows the True Positive (TP) and False Positive (FP) ratio for all possible threshold values and the Area Under Curve is a measurement of the effectiveness of the model.

Concerning regression metrics, **Mean Squared Error** (MSE) and **Mean Absolute Error** (MAE) are the most popular metrics. MSE is calculated as:

$$\frac{\sum_{i=1}^{n} y_i - \lambda(x_i)^2}{n} \tag{3}$$

where y_i is the truth value of the data-point while x_i is the predicted value, for the n number of instances, while MAE is:

$$\frac{\sum_{i=1}^{n} |(y_i - \lambda(x_i))|}{n} \tag{4}$$

with the MSE being more sensitive to outliers and larger errors, while the MAE is easily interpretable from a human point of view.

1.5.4 Loss and Gradients

Along with metrics of the networks, it is important to measure the training performances as well and it is accomplished through the loss function, a function which measures the difference between the predicted value and the actual one. Several loss functions exist, although the most used in classification tasks are the Binary, the Categorical, or the Sparse Categorical crossentropies, while Mean Absolute Error and Mean Squared Error are the most common for regression models. Binary crossentropy allows to handle with binary classification problems and its calculates as:

$$L(y, p) = -(y \log(p) + (1 - y) \log(1 - p)) \tag{5}$$

where y is the real label and p is the predicted one. When there are more than two classes, categorical crossentropy allows to use one-hot encoded labels while sparse categorical crossentropy is used with integer-labelled data. Categorical crossentropy is similar to binary one and is calculated as:

$$L(y) = -\sum_{j} y_{i,j} \log(p_{i,j}) \tag{6}$$

that is indeed an extension of the binary crossentropy summing the binary loss for each of the classes j. Concerning regression metrics, MSE and MAE are computed as the mean of the absolute or squared errors for each element of the training, defined as the difference between actual value and the predicted one. Loss function calculates the loss for each vector in the training set and all the losses are averaged using a cost function which has to be minimised using an *optimizer*, an algorithm aiming to optimise the weights of the network in order to

find the minimise the cost function. When the model is initialised, the weights are typically randomly assigned and the recalculation starts from a random point of the cost distribution matrix. It is then possible to compute the gradient of the cost function through the partial derivative and iteratively update the weights with a step size of η, known as *learning rate*. In 2014, Kingma and Ba introduced another optimizer called Adam, short for Adaptive Moment Estimation, an algorithm which adapts individual learning rate for each weight of the network [51].

1.5.5 Datasets

Training procedure starts with the setup of a dataset that is indeed divided into three subsets: a training, a validation and a test set. The training set represents the major subset and usually it contains 70-80% of available data, and is used to expose examples of the ground-truth to calculate the weights and to fit a model. The model is then evaluated using the validation set, which is a set of samples used to evaluate the training process and to fine tune the hyper-parameters. The training procedure reckons on iterations called *epochs*, which may be pre-set or the training may be programmatically interrupted when the model loss is not lowering, and when the training procedure is finished it is possible to modify some parameters until the trainer is satisfied with the results. After that phase of fine-tuning, the model is tested on a test set, a totally independent and unseen dataset. The test set is, indeed, a set composed of data which have never been exposed to the trainer nor to the network and it is used to verify the generalisation of the model on unrevealed data. It is the gold standard to asses model reliability on real world data and it is meant to provide an unbiased evaluation of the model. The application of different network architectures combined with different loss function and optimisers, led to an important variety in results and approaches in the field of Artificial Intelligence and, by reflex, in digital pathology.

1.5.6 Generative networks

Other than classical tasks in AI, one of the most hot-topic in the field of neural networks is the application of Generative Adversarial Networks (GAN), networks meant to generate data. This kind of models can be used to handle with several data types, as image synthesis for data augmentation, super-resolution or even Natural Language Processing or music generation [52]. Concerning their architecture, GANs are indeed the combination of two model, a generator and a discriminator. The generator is meant to understand the distribution of information in a ground truth to generate new images, while the discriminator is a binary classifier trying to discern fake from real images, and these two networks are competing each other [53], hence the training phase is actually a combined training of two adversarial network training to overwhelm each other, each one with its respective loss function. Among all the generative networks, Pix2Pix

network represents one of the most interesting examples, as it is easily adapt to different tasks. It has been firstly presented by Isola et al. during CVPR in 2017 and since that moment several application have been proposed, i.e. edge facade prediction [Figure 2], sketch drawing to cats representation, background removal, sketch to portrait transformation, and photo generation from sketch [54].

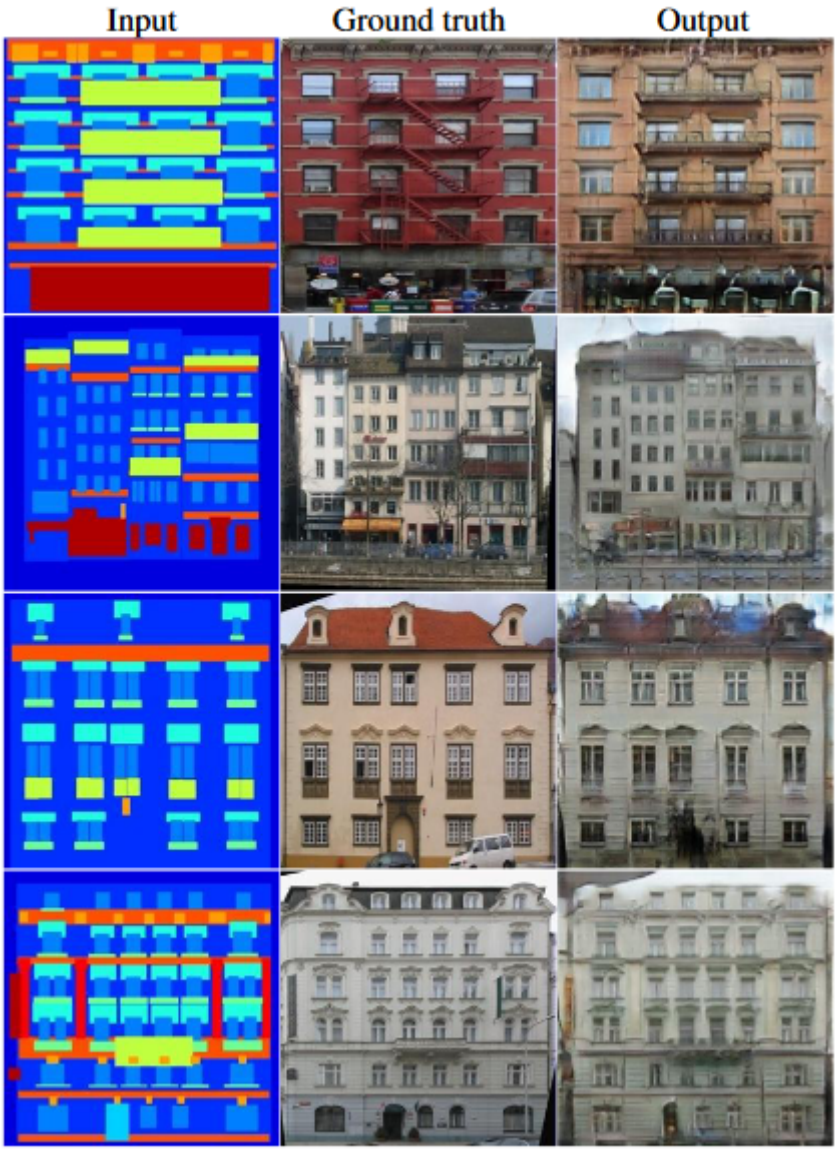

Figure 2: An example of application of Pix2Pix generative model, showing for each row the labelled image on the left, the ground truth in the middle, and the output of the model on the right. Image modified from [54]

Thus, this approach represents a valuable method to generate images even with images hugely different from the ones used in first-place, ranging from labelled images to photography, or even whole slide images with a direct application in digital pathology

1.6 Machine/Deep Learning in Digital Pathology

Artificial Intelligence in Surgical Pathology is nowadays worldwide diffused, with more and more reported results and commercial application. Most of the artificial intelligence algorithms are intended to cover a large variety of possible application in surgical pathology, e.g., grading of renal cancer [55], for prostate Gleason score classification [56], or breast cancer molecular subtyping [57] and their outcome prediction. Moreover, segmentation is another hot-topic in digital pathology, with various example of that, as the segmentation of various pathological lesions in the fields of neuropathology [58], breast cancer [59], hematopathology [60], and nephrology [61] or Oral Squamous Cell Carcinoma [62].

Moreover, Salehi and Chalechale proposed a stain normalisation method based on a Pix2Pix network, with impressive results [63] and this kind of tools may be also used to convert H&E stained slides to other staining, as described by De Haan et al. [64] and Liu et al. [65] in 2021. De Haan describes a method to convert renal H&E stained slides into three special stains, Masson's Trichrome, periodic acid-Schiff, and Jones silver stain, showing a significant increase in diagnosis accuracy when using H&E images with virtual special stains compared to H&E alone. Then, Liu describes a method to generate Ki-67 images from H&E stained slides on neuroendocrine tumour and breast cancer images using an unpaired stain transfer method and achieving a great concordance between virtual and real images.

Despite several examples of AI application to digital pathology, it is worth to note that there are some major issues to deal with when approaching to histopathology image analysis. In first instance, whole slide images are huge images, being each file in the range of gigabytes with a pixel size in the order of 100,000×100,000 pixels [27]. Hence, it is impossible to process whole slides at once, and the most common procedure involves image tiling and resizing. Consequently, there is not a standard procedure and the tiles dimension (representing the Field of View (FOV)) depends on the task, i.e. cell detection may require high magnification and small tiles, while glands detection can work with low magnification images. Eventually, tiles can be used as final input images or may need a further resizing (same FOV with a lower resolution), and questions about the best methodological approach are still open. Furthermore, WSI analysis algorithms are typically evaluated at three levels, depending on the task:

- **Tiles level**. As training is performed on tiles, its metrics are measured at tile level, as well as the validation ones.
- **WSI level**. Though the algorithms are trained to perform tasks on tiles, the clinical routine relies on WSI analysis, therefore tiles metrics are not necessarily a trustworthy measure of algorithm effectiveness. Indeed, it is important to investigate if errors in tiles classification or segmentation are unspecific or systematically happens with a specific subset of images (i.e. intensively stained slides, wrong white balancing , etc.) resulting

in misinterpreted slides.

- **Patient level**. Multiple slides are generally produced from each patient, so at the end of the day all slides assessment is combined in a single diagnosis. Is then important to verify the algorithm's effectiveness to produce a final diagnosis with the combination of all the outputs. In fact, it may result that most of the slides which belong to the same patient are misinterpreted as the can show specific characteristics (i.e. tumour grade, morphology, etc.), leading to a wrong diagnosis.

Hence, digital pathology algorithms demand for a multilevel evaluation depending on the specific purpose of each model, and the development and evaluation of the model need to be done keeping in mind the different information about the trustworthiness provided by each level of evaluation.

Another important aspect of machine and deep learning methods in digital pathology regards datasets. As mentioned beforehand, digitisation of anatomical pathology is a relatively recent event and there are few hospital with a full digital workflow, albeit increasing in last years [66], resulting a few public available datasets and in difficulties for research groups interested in the development of new algorithms. Along with the lack of datasets, there is also a lack of annotation because of the huge dimension of image and the need for highly specialised figures (pathologists) to annotate images, strongly limiting the production of AI algorithms. However, more and more results are published every year, with an exponential growth of new available algorithms, and many application have been proposed for less diffuse cancer too, as the Oral Squamous Cell Carcinoma.

1.6.1 AI applications in Oral Squamous Cell Carcinoma analysis

In a recent systematic review, Mahmood et al. showed that there is an increasing interest of AI applications to OSCC WSI analysis [67]. They reported 11 selected studies about the usage of AI-based analysis of features in oral lesions. Four papers out of 11, refer to OSCCs, namely [68]–[71], one is about oral epithelial dysplasia [72], five about oral submucous fibrosis [73]–[77], and the last one is about oropharyngeal squamous cell carcinoma [70]. Then, in 2020 we demonstrated the possible prediction of Ki-67 immunohistochemical positivity on OSCC WSI, using a machine learning approach [78].

Hence, the aim of this project is to evaluate the application of a generative network for the assessment of Ki-67, a promising biomarker, in OSCC, in continuity with our previous results.

2 Aim of the study

Ki-67 is a reliable prognostic biomarker in several human solid tumours, with a promising role as biomarker in Head and Neck Squamous Cell Carcinoma too. However, with the continuous increase in the number of immunohistochemical biomarker, it is hard to keep running the entire anatomical pathology workflow, especially in smaller or remote-areas' hospitals. Digital Pathology may represent an important helper, particularly with the application of generative networks that may allow the generation of synthetic immunostained whole slide images, decreasing the need for actual laboratory activity, with a reduction of anatomical pathology laboratories' workload. Hence, we collected 175 OSCC FFPE samples from the archive of the Anatomical Pathology unit of the University of Naples "Federico II", aiming to develop a deep learning model to convert H&E stained Whole Slide Images to anti-Ki-67 IHC stained Whole Slide Images, in order to reduce turn-around time, laboratory consumables usage and to improve patients management.

3 Materials and methods

3.1 Study Population

175 OSCC FFPE tumour samples from surgical resections, from the Pathology Unit's archive of the University of Naples "Federico II", were used to build four tissue microarrays (TMAs), and the most representative areas from each selected paraffin block were selected, at least in duplicate. TMAs were built and stained as described in [78], and a total of 349 individual cores were obtained. WSI were scanned using a Leica Aperio AT2 scanner with a 40x magnifier. After slide scanning, TMAs slide coverslips were remove by soaking the slides in xylene. The slides were then rehydrated in decreasing concentrations of ethanol and then destained using a solution of HCl 0.3% for 4 min. After destaining, the slides were rinsed in tap water and subsequently immunostained with the antibody anti-Ki-67. Immunohistochemical staining was performed on a Ventana Benchmark Ultra (Ventana Medical Systems Inc., Tucson, AZ, USA) using the rabbit monoclonal antibody anti-Ki-67 (clone 30-9, Ventana Medical Systems Inc.) in accordance with manufacturers' recommendations. The new IHC-stained slides were then digitized.

3.2 Virtual Staining

3.2.1 Dataset generation

All the 349 individual cores, for both H&E and IHC images, were dearrayed using QuPath [79] and the matching cores were manually aligned to achieve a nearly perfect pixel to pixel correspondence. After a strict quality check phase, which led to the discharge of 184 cores, the 165 remaining cores where sorted in three different datasets, training (approximately 60% of cores), test (10%), and validation (30%). Because of the nature of the experiment, the validation of the algorithm has been performed measuring the verisimilitude of the generated images and the concordance between real and synthetic images. As each core is on average 4500×4500 pixel, they have been tiled obtaining 31428, 6109, and 16560 tiles for each dataset. Validation cores were reassembled to obtain virtual cores. A brief summary of the datasets is illustrated in Table 3.

	Train	Test	Validation
Cores	95	19	51
Tiles	31428	6109	16560

Table 3: Cores and tiles distribution in train, validation, and test set. Tiles are 256×256 patches extracted from 4500×4500 cores images

3.2.2 Pix2Pix implementation

A Pix2Pix model has been trained on the training and test dataset to convert H&E images into anti-Ki-67 Immunostained images for 30 epochs using the default settings as described by Isola et al. [54]. In particular, the model consists of two part, the discriminator and the generator. During the training, the generator is trained to generate synthetic IHC images while the discriminator is trained to tell synthetic images from the actual ones. In our training, we used Adam optimiser with a learning rate of 2×10^{-4} and a momentum of 0.5 for both generator and optimiser, while the random initialisr has been set with a mean of 0 and a standard deviation of 0.02. The model has been training using two GPU Nvidia 2080 Ti. For each epoch we saved a checkpoint and at the end of the training we restored the best generative model. Validation set has been used to asses the model efficiency in two ways. In first place, we tested the ability of two pathologists to discern actual images from synthetic one on a small subset of the validation set, to evaluate the likelihood of synthetic images. Then, we counted positive cells for each core (both actual and synthetic) to evaluate the concordance between real Ki-67-positivity and the predicted one.

3.2.3 IHC Quantification

Immunopositivity has been measured using QuPath 0.2.3. After setting image type and defining stains (H&E, DAB, and Background) values, the Simple Tissue Detector has been used to define tissue areas in the images. Then, we used the "Positive cell detection" algorithm to detect and quantify positive cells within the core, and the procedure has been automated using a Groovy script [Code 1]. As IHC intensity is sometimes highly variable, the positivity threshold has been manually adjusted for some cores.

```
//QuPath Groovy Script
//Simple Tissue Detector
runPlugin('qupath.imagej.detect.tissue.SimpleTissueDetection2', '{"threshold": 220,
    "requestedDownsample": 50.0, "minAreaPixels": 100000.0, "maxHoleAreaPixels":
    500.0, "darkBackground": false, "smoothImage": true, "medianCleanup": true,
    "dilateBoundaries": false, "smoothCoordinates": true, "excludeOnBoundary": false,
    "singleAnnotation": true}');

//PositiveCellDetection
runPlugin('qupath.imagej.detect.cells.PositiveCellDetection',
    '{"detectionImageBrightfield": "Optical density sum", "backgroundRadius": 15.0,
    "medianRadius": 6.0, "sigma": 6.0, "minArea": 10.0, "maxArea": 1000.0,
    "threshold": 0.1, "maxBackground": 2.0, "watershedPostProcess": true,
    "excludeDAB": false, "cellExpansion": 0.0, "includeNuclei": true,
    "smoothBoundaries": true, "makeMeasurements": true, "thresholdCompartment":
```

```
"Nucleus: DAB OD mean", "thresholdPositive1": 0.2, "thresholdPositive2": 0.4,
"thresholdPositive3": 0.6000000000000001, "singleThreshold": true}');
```

Code 1: QuPath Groovy script to perform tissue detection and cell positivity count

3.2.4 Android application

To test the likelihood of the synthetic images, an Android application has been developed. The test is conducted showing a series of coupled IHC images, A and B, and the pathologists had to identify the synthetic one clicking on the relative button. On each trial, the pathologist were shown each image for 5 second, after which the images disappeared and unlimited time to respond as to which was fake was given. Each session consisted in a brief trial to understand the functioning of the application followed by the actual survey. The experimental setting has been adapted from the one used by Isola et al. [54]. The application is made of three activities (I) Main Activity, (II) Trial Activity, and (III) Survey Activity.

Main Activity The Main Activity is composed of three buttons, namely "Trial", "Survey", and "Settings", respectively linking to the the activities and to a setting menu. The setting menu allows to activate/deactivate the hiding of the images and its timer, along with the number of images for each session. Setting values were saved as SharedPreferences.

Trial Activity The Trial Activity is made of a disclaimer about the survey with an explanation of the modality, the actual test and a feedback about the test. The trial starts clicking the "Start" button in the disclaimer, and three images are showed to test the application. At the end of the test, a feedback is provided.

Survey Activity The Survey Activity is an exact copy of the trial one, but showing a number of images as fixed in settings. We showed thirty images to each pathologist.

The resulting application has been installed on a Android tablet and the pathologist have been supported only during the trial phase.

3.2.5 Code and Statistical Analysis

All code and statistical analysis has been performed using Python 3.9 with the packages Seaborn 0.11.1, Matplotlib 3.4.2, Pandas 1.3.0, Scikit-learn 0.241.2, OpenCv 4.5.3.56, and Tensorflow 2.6.0. The android application has been developed using Kotlin with Android Studio. All codes are available upon request.

4 Results

To fully elucidate the possible application of generative networks to OSCC as a method to predict the Ki-67 immunohistochemical positivity of H&E-stained cells, we have identified multiple project endpoints. The endpoints were:

- Dataset Generation
- Plausible Synthetic Images Generation
- Concordant Synthetic IHC and real IHC positivity

4.1 Dataset Generation

In order to generated a dataset of images of Ki-67 immunostained tumour samples perfectly matching the corresponding H&E

In first instance, we generated a dataset consisting of 349 cores stained with H&E and, after a destaining procedured, restained with IHC, they were manually aligned to create a high-quality matching-images dataset. After a dearray preprocessing, for each core the IHC image has been manually overlaid to the matching H&E and the images have been precisely aligned using Adobe Photoshop 2021. Then, the single layers (H&E and IHC) have been saved as individual images, as shown in Figure 3.

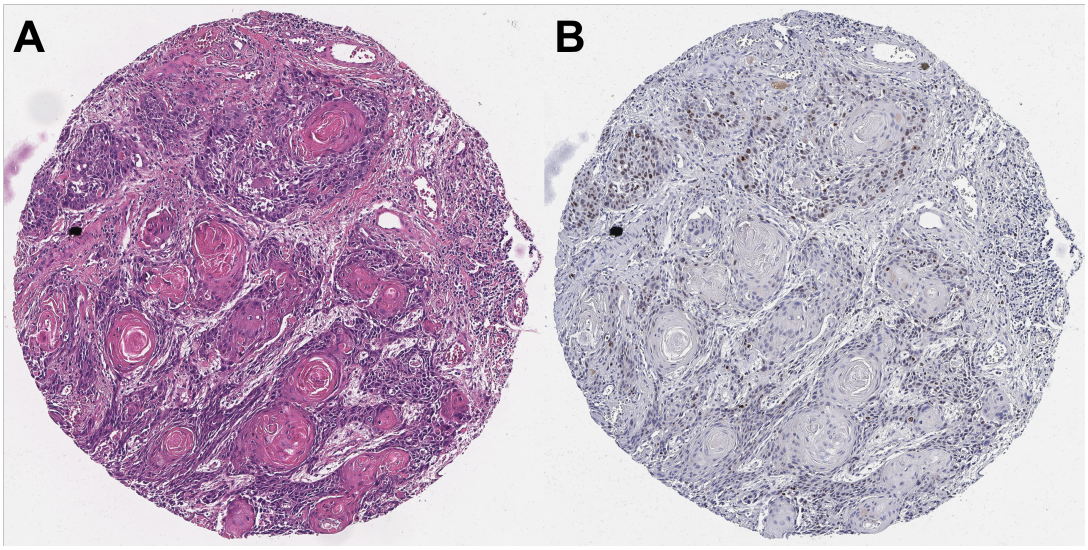

Figure 3: A representative image of two aligned cores, with the H&E in (A) and the IHC in (B). Cores have been manually annotated to a pixel level

Then, cores were subdivided in tiles and each tile was stitched together with its counterpart, in order to create the actual training dataset, with some examples shown in Figure 4.

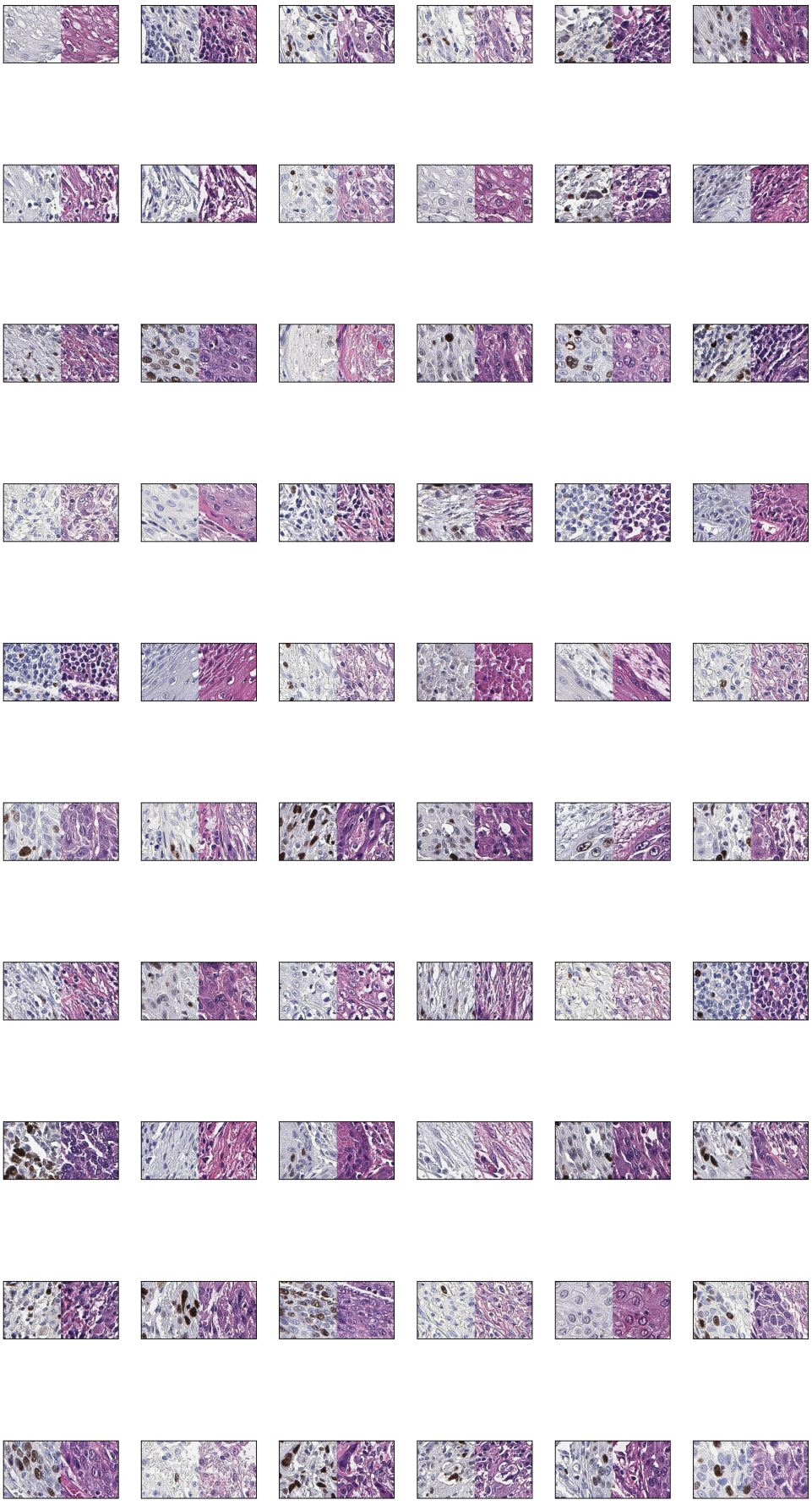

Figure 4: A grid showing some representative images of tiles. Each couple is made of target image (immunohistochemistry, left) and input image (H&E, right)

Finally, images were sorted into three datasets: training, test, and validation, as described in materials and methods and a Pix2Pix model has been trained. As shown in Figure 5, synthetic IHC images generated using the trained model appeared very similar to actual immunohistochemical images.

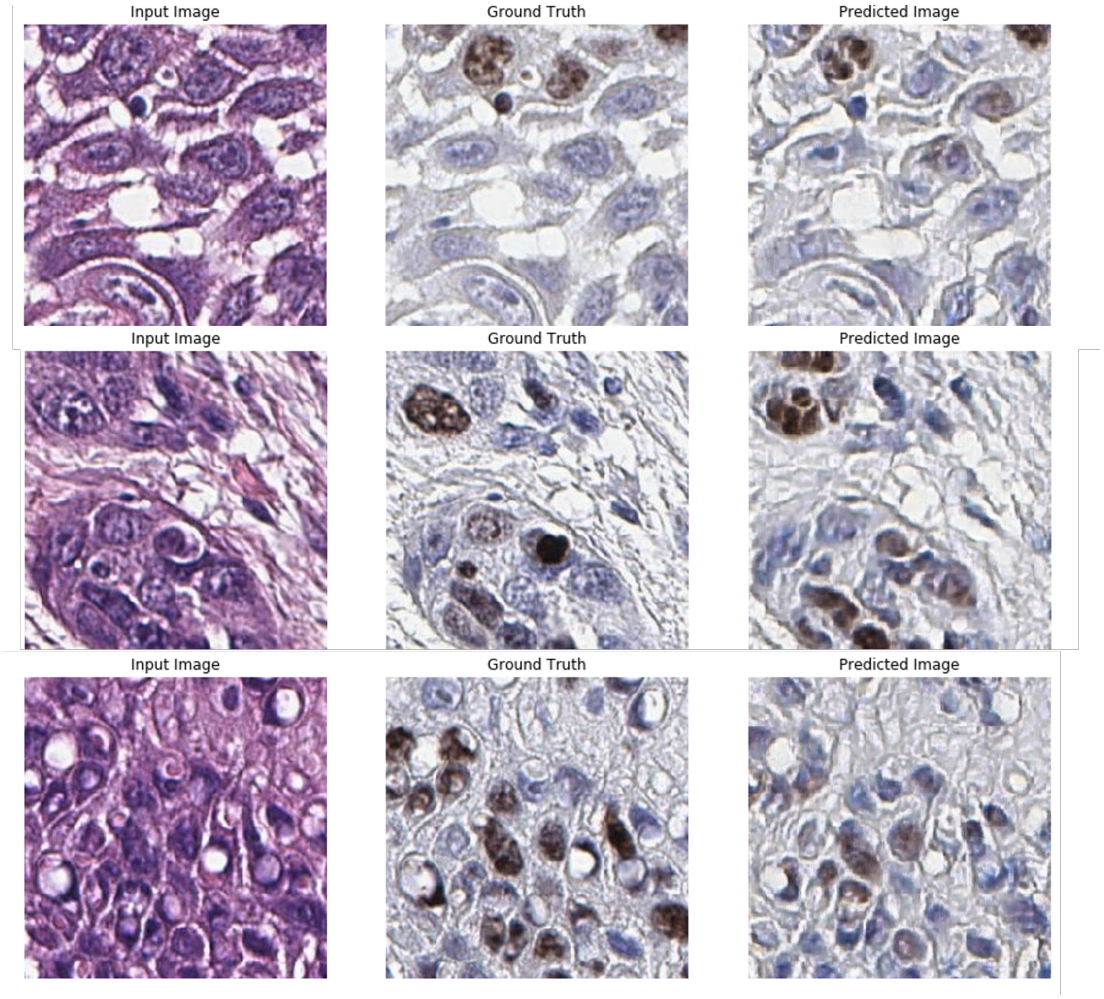

Figure 5: Three representative images of IHC prediction procedure. From left: Input H&E, Actual IHC, Synthetic IHC.

4.2 Synthetic Images Likelihood

As the proposed method is intended to generate a virtual staining, the output should appear realistic. Indeed, pathologists are used to work with actual IHC images, so realistic images (other than accurately predictions) are mandatory to assure pathologists compliance with the algorithm application. Hence, the likelihood of synthetic images has been assessed asking two pathologists to tell the synthetic image from the actual one in a side-by-side blinded comparison. The test has been performed in a dedicated android application, with a user-friendly GUI, developed to allow pathologist to evaluate model results [Figure 6].

Figure 6: Screenshots of the application developed to assess the likelihood of model's output images

Synthetic images likelihood has been measured as the percentage of synthetic images correctly identified as synthetic (True Positives), hence a percentage of 0.5 means that the pathologist were not capable to tell synthetic images from actual ones, while a value of 1 means that synthetic images are always recognisable. It resulted that among 30 images, only 17 and 16 (Pathologist 1 and Pathologist 2) were correctly recognised as synthetic, with a mean ratio of 0.55 (0.57 and 0.53), confirming the likelihood of the images and the quality of the model.

Overall, these results confirm that the images produced by the model appear realistic, assuring more compliance and confidence from pathologists. Then, we quantitatively evaluated the immunohistochemical concordance between synthetic and actual images.

4.3 Synthetic and real IHC concordance

In order to assess the concordance between synthetic and real immunohistochemical concordance, we automatically counted the number of positive cells per each core using QuPath "Positive Cell detection" as described in Materials and Methods. The comparison between synthetic and real IHC showed a moderate R^2 value (0.558), as shown in Figure 7.

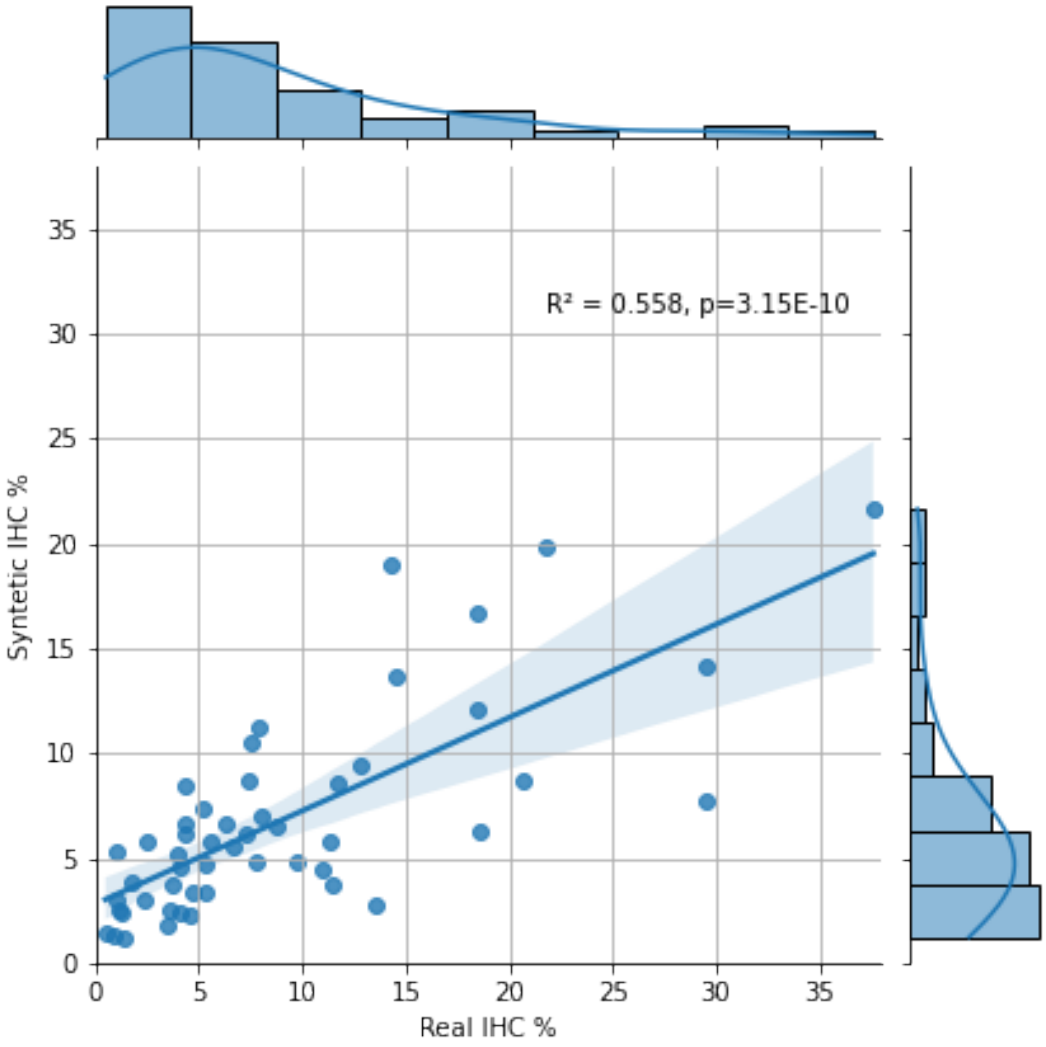

Figure 7: Jointplot showing the correlation between the predicted number of share of IHC positive cells and the actual one

Moreover, the visual inspection of the digital generated cores confirmed that the distribution pattern of positive cells matches the actual one, as showed in Figure 8.

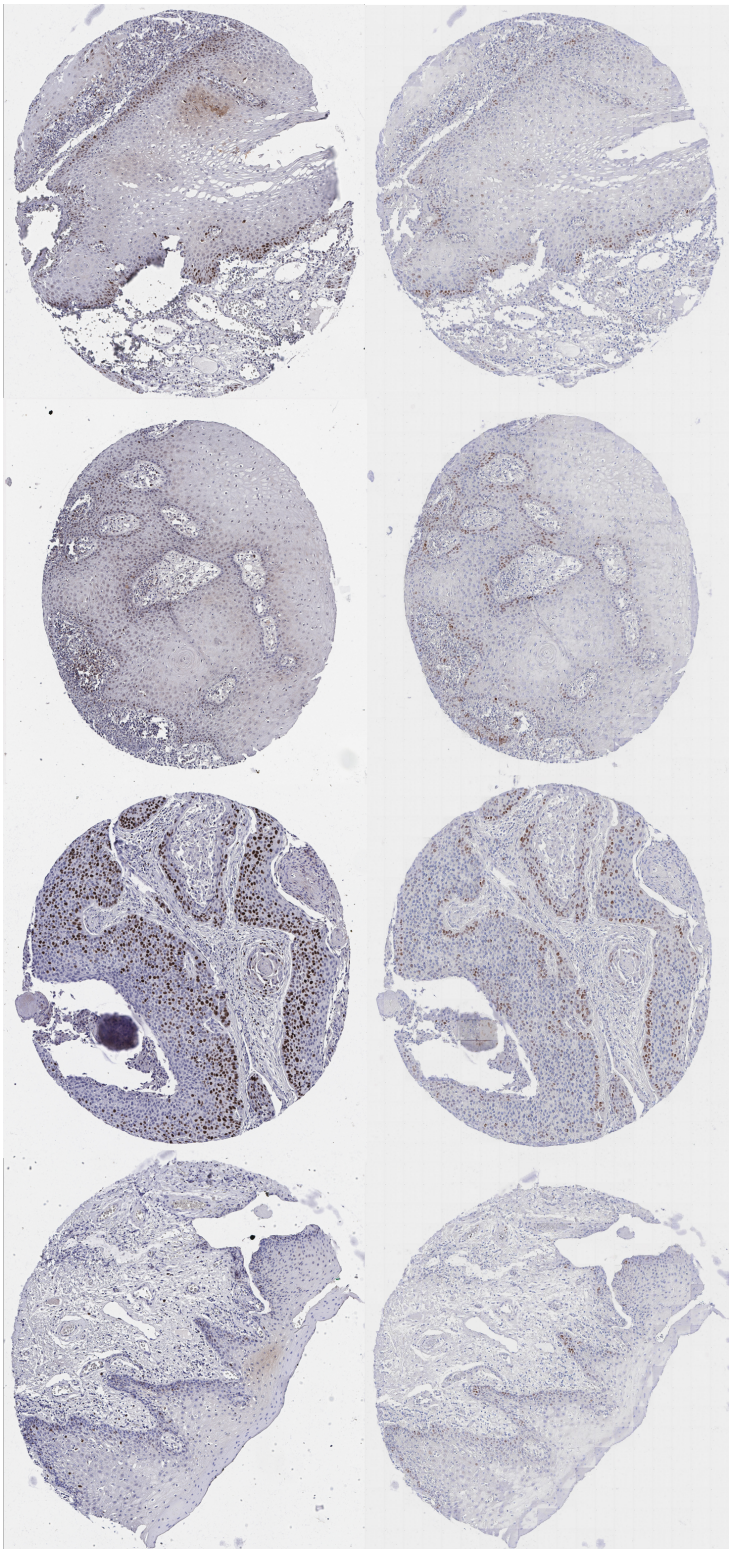

Figure 8: A comparison between actual immunohistochemistry (left) and the respective virtual one (right). It is clearly visible the close match in positivity and its pattern

Finally, categorical division in the most common used cutoff in the assessment of Ki-67

showed high accuracy rate. Indeed, when using a 5% positivity cutoff, we were able to achieve an accuracy of 74.51%, with high sensitivity (76.67%) and specificity (71.43%). With higher cutoff we had a remarkable reduction in sensitivity, although we observed a remarkable Positive Predictive Value (79.31%, 77.78%, and 75.00%).

Metric	**5%**	**10%**	**15%**
True Positives	23	7	3
False Positives	6	2	1
True Negatives	15	33	42
False Negatives	7	9	5
Sensitivity	76.67%	43.75%	37.50%
(95% CI)	(57.72% - 90.07%)	(19.75% - 70.12%)	(8.52% - 75.51%)
Specificity	71.43%	94.29%	97.67%
(95% CI)	(47.82% - 88.72%)	(80.84% - 99.30%)	(87.71% - 99.94%)
Accuracy	74.51%	78.43%	88.24%
(95% CI)	(60.37% - 85.67%)	(64.68% - 88.71%)	(76.13% - 95.56%)

Table 4: Summary of model IHC concordance metrics

Overall, although further studies are necessary to improve the algorithms and achieve a clinical applicable accuracy level, our results paved the way to a promising and reliable tool to produce immunohistochemical images directly from H&E stained slides.

5 Discussion and conclusion

During last 20 years pathology underwent through a radical transformation thanks to the introduction of the Digital Pathology, an innovative approach which merges computer science with pathology. The first computer science approaches proposed to surgical pathology regarded the automation of manual procedures as cell counts but during the course of time more and more algorithms have been developed, ranging from colour normalisation to automatic classification and segmentation of tumours. Nevertheless, one of the most outstanding application of AI to pathology is the prediction of molecular or morphological characteristics through the usage of virtual staining. The expression "virtual staining" refers to the employment of AI to obtain synthetic images of a desired target staining (as IHC, Trichrome or other staining) from a source image (as H&E), without the need of an actual laboratory activity and reducing the reagent consumption. Among the proposed algorithms, the prediction of immunohistochemistry is one of the most interesting topic because of the continuous increase in the number of tumoural biomarkers and the consequent impact on hospital budget and activity, hence we intended to develop an AI algorithm to predict the immunohistochemical positivity to Ki-67, a well-known proliferation marker related to a worse prognosis in several tumours.

We collected a series of 175 OSCC FFPE tumour samples arranged in four microarrays, and after a strict quality check phase we selected 165 individual cores that were dearrayed and tiled to create our training, validation and test set. Then, we trained a Pix2Pix model to achieve virtual stained cores from H&E images. The obtained images were assessed for their likelihood to real ones at a tile-level, while concordance between real and synthetic Ki-67 immunohistochemical expression has been assessed at a core-level with both continuous and categorical quantification through QuPath.

Virtual images deceived pathologists in 45% of cases, confirming the likelihood of the generated images. Moreover, the synthetic and real IHC positivity, measured using QuPath algorithms, showed a highly significant concordance ($R^2 = 0.56$, $p<0.001$). However, Ki-67 is assessed as low-high expression, although the cutoff value is still not defined and usually ranges between 5%-15% of positive cells. Hence, we assessed the concordance with three cutoffs (5%, 10%, and 15%), and achieved a remarkable positive predictive value of 79.31%, 77.78% and 75.00%, respectively, confirming the quality of the algorithm.

Overall, our results confirm the feasibility of a valid IHC-prediction algorithm, even with the limitation of our study. Indeed, we performed our training on TMA cores, which were made with samples collected in different time period and then with an intrinsic variability in IHC staining. Moreover, we used a relative small amount of OSCC cores, all coming from our institute, so further studies are necessary to assess the generalizability of our algorithm. In conclusion, although larger and wider studies are necessary to fine tune and improve our algorithm, the results we achieved represent a promising starting point to further develop a virtual staining

protocol to reduce the turn-around time and the material and reagents consumption, in order to achieve an actual *digital* pathology. Virtual staining protocols may result, in the future, in a faster and economical way to gather information about protein expression and mutational status of patients, improving the effectiveness of therapies and also in hospitals in developing countries as they usually cannot access to immuno-stainers, reagent supplies and adequate trained stuff, enabling high-level patient-care requiring only a slide scanner.

List of Figures

1 Neural Network Scheme 11
2 Example of Pix2Pix application 15
3 Example of aligned cores 22
4 Dataset representation 23
5 Example of prediction 24
6 Android application 25
7 Synthetic and real immunopositivity jointplot 27
8 Prediction comparison 28

List of Tables

1 Glossary of terms 8
2 Metrics case study 12
3 Dataset composition 19
4 IHC Analysis metrics 29

References

[1] Manuel Salto-Tellez, Perry Maxwell, and Peter Hamilton. "Artificial intelligence-the third revolution in pathology". en. In: *Histopathology* 74.3 (Feb. 2019), pp. 372–376. ISSN: 03090167. DOI: 10.1111/his.13760. URL: https://onlinelibrary.wiley.com/doi/10.1111/his.13760 (visited on 10/01/2021).

[2] Anthony A. Gal. "In Search of the Origins of Modern Surgical Pathology". en-US. In: *Advances in Anatomic Pathology* 8.1 (Jan. 2001), pp. 1–13. ISSN: 1072-4109. URL: https://journals.lww.com/anatomicpathology/Fulltext/2001/01000/In_Search_of_the_Origins_of_Modern_Surgical.1.aspx (visited on 10/01/2021).

[3] Ludwig A. Sternberger, Paul H. Hardy, John J. Cuculis, et al. "The unlabeled antibody enzyme method of immunohistochemistry preparation and properties of soluble antigen-antibody complex (horseradish peroxidase-antihorseradish peroxidase) and its use in identification of spirochetes". en. In: *Journal of Histochemistry & Cytochemistry* 18.5 (May 1970), pp. 315–333. ISSN: 0022-1554, 1551-5044. DOI: 10.1177/18.5.315. URL: http://journals.sagepub.com/doi/10.1177/18.5.315 (visited on 10/01/2021).

[4] John R. Johnson, Peter Bross, Martin Cohen, et al. "Approval Summary: Imatinib Mesylate Capsules for Treatment of Adult Patients with Newly Diagnosed Philadelphia Chromosome-positive Chronic Myelogenous Leukemia in Chronic Phase". en. In: *Clinical Cancer Research* 9.6 (June 2003), pp. 1972–1979. ISSN: 1078-0432, 1557-3265. URL: https://clincancerres.aacrjournals.org/content/9/6/1972 (visited on 10/01/2021).

[5] J Louise Jones, Karin A Oien, Jessica L Lee, et al. "Morphomolecular pathology: setting the framework for a new generation of pathologists". en. In: *British Journal of Cancer* 117.11 (Nov. 2017), pp. 1581–1582. ISSN: 0007-0920, 1532-1827. DOI: 10.1038/bjc.2017.340. URL: http://www.nature.com/articles/bjc2017340 (visited on 10/01/2021).

[6] R. Ferreira, B. Moon, J. Humphries, et al. "The Virtual Microscope". eng. In: *Proceedings: a conference of the American Medical Informatics Association. AMIA Fall Symposium* (1997), pp. 449–453. ISSN: 1091-8280.

[7] James Bacus. "Method and apparatus for acquiring and reconstructing magnified specimen images from a computer-controlled microscope". en. US20010050999A1. Dec. 2001. URL: https://patents.google.com/patent/US20010050999/en (visited on 10/01/2021).

[8] Andrew J. Evans, Runjan Chetty, Blaise A. Clarke, et al. "Primary frozen section diagnosis by robotic microscopy and virtual slide telepathology: the University Health Network experience". eng. In: *Human Pathology* 40.8 (Aug. 2009), pp. 1070–1081. ISSN: 1532-8392. DOI: 10.1016/j.humpath.2009.04.012.

[9] Leslie A. Bruch, Barry R. De Young, Clarence D. Kreiter, et al. "Competency assessment of residents in surgical pathology using virtual microscopy". eng. In: *Human Pathology* 40.8 (Aug. 2009), pp. 1122–1128. ISSN: 1532-8392. DOI: 10.1016/j.humpath.2009.04.009.

[10] Mike Isaacs, Jochen K. Lennerz, Stacey Yates, et al. "Implementation of whole slide imaging in surgical pathology: A value added approach". In: *Journal of Pathology Informatics* 2 (Aug. 2011), p. 39. ISSN: 2153-3539. DOI: 10.4103/2153-3539.84232. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3169923/ (visited on 10/01/2021).

[11] Andrew J. Evans, Mohamed E. Salama, Walter H. Henricks, et al. "Implementation of Whole Slide Imaging for Clinical Purposes: Issues to Consider From the Perspective of Early Adopters". eng. In: *Archives of Pathology & Laboratory Medicine* 141.7 (July 2017), pp. 944–959. ISSN: 1543-2165. DOI: 10.5858/arpa.2016-0074-OA.

[12] Nikolas Stathonikos, Mitko Veta, André Huisman, et al. "Going fully digital: Perspective of a Dutch academic pathology lab". eng. In: *Journal of Pathology Informatics* 4 (2013), p. 15. ISSN: 2229-5089. DOI: 10.4103/2153-3539.114206.

[13] Huazhang Guo, Joe Birsa, Navid Farahani, et al. "Digital pathology and anatomic pathology laboratory information system integration to support digital pathology sign-out". In: *Journal of Pathology Informatics* 7 (May 2016), p. 23. ISSN: 2229-5089. DOI: 10.4103/2153-3539.181767. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4872480/ (visited on 10/01/2021).

[14] Filippo Fraggetta, Salvatore Garozzo, GianFranco Zannoni, et al. "Routine digital pathology workflow: The Catania experience". en. In: *Journal of Pathology Informatics* 8.1 (2017), p. 51. ISSN: 2153-3539. DOI: 10.4103/jpi.jpi_58_17. URL: http://www.jpathinformatics.org/text.asp?2017/8/1/51/221132 (visited on 07/30/2021).

[15] Daniel E. Johnson, Barbara Burtness, C. René Leemans, et al. "Head and neck squamous cell carcinoma". In: *Nature Reviews Disease Primers* 6.1 (Nov. 26, 2020), p. 92. ISSN: 2056-676X. DOI: 10.1038/s41572-020-00224-3. URL: https://doi.org/10.1038/s41572-020-00224-3.

[16] Nadarajah Vigneswaran and Michelle D. Williams. "Epidemiological Trends in Head and Neck Cancer and Aids in Diagnosis". In: *Oral and maxillofacial surgery clinics of North America* 26.2 (May 2014), pp. 123–141. ISSN: 1042-3699. DOI: 10.1016/j.coms.2014.01.001. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4040236/ (visited on 04/19/2021).

[17] Saman Warnakulasuriya and John S. Greenspan. "Epidemiology of Oral and Oropharyngeal Cancers". In: *Textbook of Oral Cancer: Prevention, Diagnosis and Management*. Ed. by Saman Warnakulasuriya and John S. Greenspan. Textbooks in Contemporary Dentistry. Cham: Springer International Publishing, 2020, pp. 5–21. ISBN: 978-3-030-32316-5.

DOI: 10.1007/978-3-030-32316-5_2. URL: https://doi.org/10.1007/978-3-030-32316-5_2 (visited on 04/19/2021).

[18] Global Burden of Disease Cancer Collaboration, Christina Fitzmaurice, Degu Abate, et al. "Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-Years for 29 Cancer Groups, 1990 to 2017: A Systematic Analysis for the Global Burden of Disease Study". In: *JAMA oncology* 5.12 (Dec. 1, 2019), pp. 1749–1768. ISSN: 2374-2445. DOI: 10.1001/jamaoncol.2019.2996.

[19] Ismail Yazdi Nasim Taghavi. "Prognostic Factors of Survival Rate in Oral Squamous Cell Carcinoma: Clinical, Histologic, Genetic and Molecular Concepts". In: (), p. 6.

[20] H. B. Santos, T. K. dos Santos, A. R. Paz, et al. "Clinical findings and risk factors to oral squamous cell carcinoma in young patients: A 12-year retrospective analysis". In: *Med Oral Patol Oral Cir Bucal* 21.2 (2016), e151–156.

[21] N. Sathish, X. Wang, and Y. Yuan. "Human Papillomavirus (HPV)-associated Oral Cancers and Treatment Strategies". In: *Journal of Dental Research* 93.7 (July 1, 2014). Publisher: SAGE Publications Inc, 29S–36S. ISSN: 0022-0345. DOI: 10.1177/0022034514527969. URL: https://doi.org/10.1177/0022034514527969 (visited on 04/26/2021).

[22] Brian O'Sullivan, James Brierley, David Byrd, et al. "The TNM classification of malignant tumours—towards common understanding and reasonable expectations". In: *The Lancet. Oncology* 18.7 (July 2017), pp. 849–851. ISSN: 1470-2045. DOI: 10.1016/S1470-2045(17)30438-2. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5851445/ (visited on 04/26/2021).

[23] Leon J. Wils, Jos B. Poell, Ilkay Evren, et al. "Incorporation of differentiated dysplasia improves prediction of oral leukoplakia at increased risk of malignant progression". In: *Modern Pathology* 33.6 (June 2020). Number: 6 Publisher: Nature Publishing Group, pp. 1033–1040. ISSN: 1530-0285. DOI: 10.1038/s41379-019-0444-0. URL: https://www.nature.com/articles/s41379-019-0444-0 (visited on 04/29/2021).

[24] Séamus S. Napier and Paul M. Speight. "Natural history of potentially malignant oral lesions and conditions: an overview of the literature". In: *Journal of Oral Pathology & Medicine* 37.1 (2008). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1600-0714.2007.00579.x, pp. 1–10. ISSN: 1600-0714. DOI: https://doi.org/10.1111/j.1600-0714.2007.00579.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-0714.2007.00579.x (visited on 04/29/2021).

[25] Sjoukje F. Oosting and Robert I. Haddad. "Best Practice in Systemic Therapy for Head and Neck Squamous Cell Carcinoma". In: *Frontiers in Oncology* 9 (Aug. 27, 2019). ISSN:

2234-943X. DOI: 10.3389/fonc.2019.00815. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6718707/ (visited on 05/04/2021).

[26] *Glossary of Terms*. en. URL: https://digitalpathologyassociation.org/glossary-of-terms_1 (visited on 07/30/2021).

[27] Fusheng Wang, Tae W. Oh, Cristobal Vergara-Niedermayr, et al. "Managing and Querying Whole Slide Images". In: *Proceedings of SPIE* 8319 (Feb. 2012), 83190J. ISSN: 1996-756X. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3405921/ (visited on 09/30/2021).

[28] Toby C. Cornish, Ryan E. Swapp, and Keith J. Kaplan. "Whole-slide imaging: routine pathologic diagnosis". eng. In: *Advances in Anatomic Pathology* 19.3 (May 2012), pp. 152–159. ISSN: 1533-4031. DOI: 10.1097/PAP.0b013e318253459e.

[29] Sanjay Mukhopadhyay, Michael D. Feldman, Esther Abels, et al. "Whole Slide Imaging Versus Microscopy for Primary Diagnosis in Surgical Pathology: A Multicenter Blinded Randomized Noninferiority Study of 1992 Cases (Pivotal Study)". en-US. In: *The American Journal of Surgical Pathology* 42.1 (Jan. 2018), pp. 39–52. ISSN: 0147-5185. DOI: 10.1097/PAS.0000000000000948. URL: https://journals.lww.com/ajsp/Fulltext/2018/01000/Whole_Slide_Imaging_Versus_Microscopy_for_Primary.6.aspx (visited on 07/30/2021).

[30] Elizabeth A. Krupinski, Achyut K. Bhattacharyya, and Ronald S. Weinstein. "Telepathology and Digital Pathology Research". en. In: (2016). Ed. by Keith J. Kaplan and Luigi K.F. Rao, pp. 41–54. DOI: 10.1007/978-3-319-20379-9_5. URL: https://doi.org/10.1007/978-3-319-20379-9_5 (visited on 08/02/2021).

[31] Miao Cui and David Y. Zhang. "Artificial intelligence and computational pathology". en. In: *Laboratory Investigation* 101.4 (Apr. 2021), pp. 412–422. ISSN: 1530-0307. DOI: 10.1038/s41374-020-00514-0. URL: https://www.nature.com/articles/s41374-020-00514-0 (visited on 08/02/2021).

[32] Amir Vudka. "The Golem in the age of artificial intelligence". en. In: (July 2020). Publisher: Amsterdam University Press. ISSN: 2213-0217. DOI: 10.25969/MEDIAREP/14326. URL: https://mediarep.org/handle/doc/15295 (visited on 08/02/2021).

[33] "Meet the 18th Century chess machine". en-GB. In: *BBC News* (). URL: https://www.bbc.com/news/av/magazine-21882456 (visited on 08/02/2021).

[34] Stephen Muggleton. "Alan Turing and the development of Artificial Intelligence". en. In: *AI Communications* 27.1 (Jan. 2014). Publisher: IOS Press, pp. 3–10. ISSN: 0921-7126. DOI: 10.3233/AIC-130579. URL: https://content.iospress.com/articles/ai-communications/aic579 (visited on 08/02/2021).

[35] Snehashish Chakraverty, Deepti Moyi Sahoo, and Nisha Rani Mahato. “McCulloch–Pitts Neural Network Model”. en. In: (2019). Ed. by Snehashish Chakraverty, Deepti Moyi Sahoo, and Nisha Rani Mahato, pp. 167–173. DOI: 10.1007/978-981-13-7430-2_11. URL: https://doi.org/10.1007/978-981-13-7430-2_11 (visited on 08/02/2021).

[36] Patrick Jungwirth, David Richie, and Barry Secrest. “Analog Convolutional Neural Network”. In: 2 (Mar. 2020). ISSN: 1558-058X, pp. 1–6. DOI: 10.1109/SoutheastCon44009.2020.9368273.

[37] John McCarthy, Marvin L. Minsky, Nathaniel Rochester, et al. “A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955”. en. In: *AI Magazine* 27.4 (Dec. 2006). Number: 4, pp. 12–12. ISSN: 2371-9621. DOI: 10.1609/aimag.v27i4.1904. URL: https://ojs.aaai.org/index.php/aimagazine/article/view/1904 (visited on 08/02/2021).

[38] Kai Arulkumaran, Marc Peter Deisenroth, Miles Brundage, et al. “Deep Reinforcement Learning: A Brief Survey”. In: *IEEE Signal Processing Magazine* 34.6 (Nov. 2017). Conference Name: IEEE Signal Processing Magazine, pp. 26–38. ISSN: 1558-0792. DOI: 10.1109/MSP.2017.2743240.

[39] Ajay Shrestha and Ausif Mahmood. “Review of Deep Learning Algorithms and Architectures”. In: *IEEE Access* 7 (2019). Conference Name: IEEE Access, pp. 53040–53065. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2019.2912200.

[40] D. H. Hubel and T. N. Wiesel. “Receptive fields of single neurones in the cat’s striate cortex”. en. In: *J. Physiol.* 148.574-591 (1959), p. 18.

[41] Russell A. Kirsch. “Computer determination of the constituent structure of biological images”. en. In: *Computers and Biomedical Research* 4.3 (June 1971), pp. 315–328. ISSN: 0010-4809. DOI: 10.1016/0010-4809(71)90034-6. URL: https://www.sciencedirect.com/science/article/pii/0010480971900346 (visited on 09/08/2021).

[42] T S Huang. “Computer Vision: Evolution and Promise”. en. In: *CERN School of Computing* (1996), p. 5.

[43] S B Kotsiantis. “Supervised Machine Learning: A Review of Classification Techniques”. en. In: (), p. 20.

[44] Yash Agarwal and G. Poornalatha. “Analysis of the Nearest Neighbor Classifiers: A Review”. en. In: Advances in Intelligent Systems and Computing (2021). Ed. by Niranjan N. Chiplunkar and Takanori Fukao, pp. 559–570. DOI: 10.1007/978-981-15-3514-7_43.

[45] S. B. Kotsiantis. “Decision trees: a recent overview”. en. In: *Artificial Intelligence Review* 39.4 (Apr. 2013), pp. 261–283. ISSN: 1573-7462. DOI: 10.1007/s10462-011-9272-4. URL: https://doi.org/10.1007/s10462-011-9272-4 (visited on 08/03/2021).

[46] Khaled Fawagreh, Mohamed Medhat Gaber, and Eyad Elyan. "Random forests: from early developments to recent advancements". In: *Systems Science & Control Engineering* 2.1 (Dec. 2014), pp. 602–609. ISSN: null. DOI: 10.1080/21642583.2014.956265. URL: https://doi.org/10.1080/21642583.2014.956265 (visited on 08/03/2021).

[47] Vladimir Vapnik. *The Nature of Statistical Learning Theory*. en. 2nd ed. Information Science and Statistics. New York: Springer-Verlag, 2000. ISBN: 978-0-387-98780-4. DOI: 10.1007/978-1-4757-3264-1. URL: https://www.springer.com/gp/book/9780387987804 (visited on 08/03/2021).

[48] Mostafa Langarizadeh and Fateme Moghbeli. "Applying Naive Bayesian Networks to Disease Prediction: a Systematic Review". In: *Acta Informatica Medica* 24.5 (Oct. 2016), pp. 364–369. ISSN: 0353-8109. DOI: 10.5455/aim.2016.24.364-369. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5203736/ (visited on 08/03/2021).

[49] Wesam Ashour Barbakh, Ying Wu, and Colin Fyfe. *Review of Clustering Algorithms*. en. Ed. by Wesam Ashour Barbakh, Ying Wu, and Colin Fyfe. Studies in Computational Intelligence. Berlin, Heidelberg: Springer, 2009, pp. 7–28. ISBN: 978-3-642-04005-4. DOI: 10.1007/978-3-642-04005-4_2. URL: https://doi.org/10.1007/978-3-642-04005-4_2 (visited on 09/08/2021).

[50] Juergen Schmidhuber. "Deep Learning in Neural Networks: An Overview". en. In: *Neural Networks* 61 (Jan. 2015). arXiv: 1404.7828, pp. 85–117. ISSN: 08936080. DOI: 10.1016/j.neunet.2014.09.003. URL: http://arxiv.org/abs/1404.7828 (visited on 09/08/2021).

[51] Diederik P. Kingma and Jimmy Ba. "Adam: A Method for Stochastic Optimization". In: *arXiv:1412.6980 [cs]* (Jan. 2017). arXiv: 1412.6980. URL: http://arxiv.org/abs/1412.6980 (visited on 09/08/2021).

[52] Jie Gui, Zhenan Sun, Yonggang Wen, et al. "A Review on Generative Adversarial Networks: Algorithms, Theory, and Applications". en. In: *arXiv:2001.06937 [cs, stat]* (Jan. 2020). arXiv: 2001.06937. URL: http://arxiv.org/abs/2001.06937 (visited on 09/15/2021).

[53] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, et al. "Generative Adversarial Nets". In: *Advances in Neural Information Processing Systems*. Ed. by Z. Ghahramani, M. Welling, C. Cortes, et al. Vol. 27. Curran Associates, Inc., 2014. URL: https://proceedings.neurips.cc/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf.

[54] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, et al. "Image-to-Image Translation with Conditional Adversarial Networks". In: (2018). arXiv: 1611.07004 [cs.CV].

[55] Katherine Tian, Christopher A. Rubadue, Douglas I. Lin, et al. "Automated clear cell renal carcinoma grade classification with prognostic significance". In: *PLoS ONE* 14 (10 Oct. 2019). ISSN: 19326203. DOI: 10.1371/journal.pone.0222641. URL: https://pubmed.ncbi.nlm.nih.gov/31581201/.

[56] Hongming Xu, Sunho Park, and Tae Hyun Hwang. "Computerized Classification of Prostate Cancer Gleason Scores from Whole Slide Images". In: *IEEE/ACM Transactions on Computational Biology and Bioinformatics* (Sept. 2019), pp. 1–1. ISSN: 1545-5963. DOI: 10.1109/tcbb.2019.2941195. URL: https://pubmed.ncbi.nlm.nih.gov/31536012/.

[57] Mustafa I. Jaber, Bing Song, Clive Taylor, et al. "A deep learning image-based intrinsic molecular subtype classifier of breast tumors reveals tumor heterogeneity that may affect survival". In: *Breast Cancer Research* 22 (1 Jan. 2020). ISSN: 1465542X. DOI: 10.1186/s13058-020-1248-3. URL: https://pubmed.ncbi.nlm.nih.gov/31992350/.

[58] Ziqi Tang, Kangway V. Chuang, Charles DeCarli, et al. "Interpretable classification of Alzheimer's disease pathologies with a convolutional neural network pipeline". In: *Nature Communications* 10 (1 Dec. 2019). ISSN: 20411723. DOI: 10.1038/s41467-019-10212-1. URL: https://pubmed.ncbi.nlm.nih.gov/31092819/.

[59] Zichao Guo, Hong Liu, Haomiao Ni, et al. "A Fast and Refined Cancer Regions Segmentation Framework in Whole-slide Breast Pathological Images". In: *Scientific Reports* 9 (1 Dec. 2019). ISSN: 20452322. DOI: 10.1038/s41598-018-37492-9. URL: https://pubmed.ncbi.nlm.nih.gov/30696894/.

[60] Frederik Skou Nielsen, Mads Jozwiak Pedersen, Mathias Vassard Olsen, et al. "Automatic Bone Marrow Cellularity Estimation in H&E Stained Whole Slide Images". In: *Cytometry Part A* 95 (10 Oct. 2019), pp. 1066–1074. ISSN: 15524930. DOI: 10.1002/cyto.a.23885. URL: https://pubmed.ncbi.nlm.nih.gov/31490627/.

[61] Gloria Bueno, M. Milagro Fernandez-Carrobles, Lucia Gonzalez-Lopez, et al. "Glomerulosclerosis identification in whole slide images using semantic segmentation". In: *Computer Methods and Programs in Biomedicine* 184 (Feb. 2020). ISSN: 18727565. DOI: 10.1016/j.cmpb.2019.105273. URL: https://pubmed.ncbi.nlm.nih.gov/31891905/.

[62] Francesco Martino, Domenico D. Bloisi, Andrea Pennisi, et al. "Deep Learning-Based Pixel-Wise Lesion Segmentation on Oral Squamous Cell Carcinoma Images". en. In: *Applied Sciences* 10.22 (Jan. 2020). Number: 22 Publisher: Multidisciplinary Digital Publishing Institute, p. 8285. DOI: 10.3390/app10228285. URL: https://www.mdpi.com/2076-3417/10/22/8285 (visited on 09/15/2021).

[63] Pegah Salehi and Abdolah Chalechale. "Pix2Pix-based Stain-to-Stain Translation: A Solution for Robust Stain Normalization in Histopathology Images Analysis". In: (2020). arXiv: 2002.00647 [eess.IV].

[64] Kevin de Haan, Yijie Zhang, Jonathan E. Zuckerman, et al. "Deep learning-based transformation of H&E stained tissues into special stains". en. In: *Nature Communications* 12.1 (Aug. 2021), p. 4884. ISSN: 2041-1723. DOI: 10.1038/s41467-021-25221-2. URL: https://www.nature.com/articles/s41467-021-25221-2 (visited on 09/30/2021).

[65] Shuting Liu, Baochang Zhang, Yiqing Liu, et al. "Unpaired Stain Transfer Using Pathology-Consistent Constrained Generative Adversarial Networks". eng. In: *IEEE transactions on medical imaging* 40.8 (Aug. 2021), pp. 1977–1989. ISSN: 1558-254X. DOI: 10.1109/TMI.2021.3069874.

[66] Samar Betmouni. "Diagnostic digital pathology implementation: Learning from the digital health experience". In: *Digital Health* 7 (June 2021), p. 20552076211020240. ISSN: 2055-2076. DOI: 10.1177/20552076211020240. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8216403/ (visited on 09/30/2021).

[67] H. Mahmood, M. Shaban, B. I. Indave, et al. "Use of artificial intelligence in diagnosis of head and neck precancerous and cancerous lesions: A systematic review". In: *Oral Oncology* 110 (Nov. 2020), p. 104885. ISSN: 18790593. DOI: 10.1016/j.oraloncology.2020.104885.

[68] Yung-Nien Sun, Yi-Ying Wang, Shao-Chien Chang, et al. "Color-based tumor tissue segmentation for the automated estimation of oral cancer parameters". In: *Microscopy Research and Technique* 73 (1 Jan. 2009), NA–NA. ISSN: 1059910X. DOI: 10.1002/jemt.20746. URL: http://doi.wiley.com/10.1002/jemt.20746.

[69] T. Y. Rahman, L. B. Mahanta, C. Chakraborty, et al. "Textural pattern classification for oral squamous cell carcinoma". In: *Journal of Microscopy* 269 (1 Jan. 2018), pp. 85–93. ISSN: 13652818. DOI: 10.1111/jmi.12611. URL: https://pubmed.ncbi.nlm.nih.gov/28768053/.

[70] Shereen Fouad, David Randell, Antony Galton, et al. "Unsupervised morphological segmentation of tissue compartments in histopathological images". In: *PLOS ONE* 12 (11 Nov. 2017). Ed. by Dalin Tang, e0188717. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0188717. URL: https://dx.plos.org/10.1371/journal.pone.0188717.

[71] Dev Kumar Das, Surajit Bose, Asok Kumar Maiti, et al. "Automatic identification of clinically relevant regions from oral tissue histological images for oral squamous cell carcinoma diagnosis". In: *Tissue and Cell* 53 (Aug. 2018), pp. 111–119. ISSN: 15323072. DOI: 10.1016/j.tice.2018.06.004.

[72] Jonathan Baik, Qian Ye, Lewei Zhang, et al. "Automated classification of oral premalignant lesions using image cytometry and Random Forests-based algorithms". In: *Cellular Oncology* 37 (3 2014). ISSN: 22113436. DOI: 10.1007/s13402-014-0172-x.

[73] M. Muthu Rama Krishnan, Vikram Venkatraghavan, U. Rajendra Acharya, et al. "Automated oral cancer identification using histopathological images: A hybrid feature extraction paradigm". In: *Micron* 43 (2-3 2012). ISSN: 09684328. DOI: 10.1016/j.micron.2011.09.016.

[74] M. Muthu Rama Krishnan, Pratik Shah, Chandan Chakraborty, et al. "Statistical analysis of textural features for improved classification of oral histopathological images". In: *Journal of Medical Systems* 36 (2 Apr. 2012), pp. 865–881. ISSN: 01485598. DOI: 10.1007/s10916-010-9550-8.

[75] M. Muthu Rama Krishnan, Anirudh Choudhary, Chandan Chakraborty, et al. "Texture based segmentation of epithelial layer from oral histological images". In: *Micron* 42 (6 2011). ISSN: 09684328. DOI: 10.1016/j.micron.2011.03.003.

[76] M. Muthu Rama Krishnan, Mousumi Pal, Suneel K. Bomminayuni, et al. "Automated classification of cells in sub-epithelial connective tissue of oral sub-mucous fibrosis-An SVM based approach". In: *Computers in Biology and Medicine* 39 (12 Dec. 2009), pp. 1096–1104. ISSN: 00104825. DOI: 10.1016/j.compbiomed.2009.09.004.

[77] Muthu Rama Krishnan Mookiah, Pratik Shah, Chandan Chakraborty, et al. "Brownian motion curve-based textural classification and its application in cancer diagnosis". In: *Analytical and Quantitative Cytology and Histology* 33 (3 June 2011), pp. 158–168. ISSN: 08846812.

[78] Francesco Martino, Silvia Varricchio, Daniela Russo, et al. "A machine-learning approach for the assessment of the proliferative compartment of solid tumors on hematoxylin-eosin-stained sections". In: *Cancers* 12 (5 2020). ISSN: 20726694. DOI: 10.3390/cancers12051344.

[79] Peter Bankhead, Maurice B. Loughrey, José A. Fernández, et al. "QuPath: Open source software for digital pathology image analysis". en. In: *Scientific Reports* 7.1 (Dec. 2017), p. 16878. ISSN: 2045-2322. DOI: 10.1038/s41598-017-17204-5. URL: https://www.nature.com/articles/s41598-017-17204-5 (visited on 08/31/2021).