Digital image analysis in breast pathology—from image processing techniques to artificial intelligence

STEPHANIE ROBERTSON, HOSSEIN AZIZPOUR, KEVIN SMITH, and JOHAN HARTMAN

STOCKHOLM, SWEDEN

Breast cancer is the most common malignant disease in women worldwide. In recent decades, earlier diagnosis and better adjuvant therapy have substantially improved patient outcome. Diagnosis by histopathology has proven to be instrumental to guide breast cancer treatment, but new challenges have emerged as our increasing understanding of cancer over the years has revealed its complex nature. As patient demand for personalized breast cancer therapy grows, we face an urgent need for more precise biomarker assessment and more accurate histopathologic breast cancer diagnosis to make better therapy decisions. The digitization of pathology data has opened the door to faster, more reproducible, and more precise diagnoses through computerized image analysis. Software to assist diagnostic breast pathology through image processing techniques have been around for years. But recent breakthroughs in artificial intelligence (AI) promise to fundamentally change the way we detect and treat breast cancer in the near future. Machine learning, a subfield of AI that applies statistical methods to learn from data, has seen an explosion of interest in recent years because of its ability to recognize patterns in data with less need for human instruction. One technique in particular, known as deep learning, has produced groundbreaking results in many important problems including image classification and speech recognition. In this review, we will cover the use of AI and deep learning in diagnostic breast pathology, and other recent developments in digital image analysis. (Translational Research 2018;194:19–35)

Abbreviations: DIA = digital image analysis; AI = artificial intelligence; H&E = hematoxylin and eosin; IHC = immunohistochemistry; ISH = in situ hybridization; DCIS = ductal carcinoma in situ; ER = estrogen receptor α; PR = progesterone receptor; HER2 = human epidermal growth factor receptor 2; DNA = deoxyribonucleic acid; mRNA = messenger ribonucleic acid; WSI = whole-slide imaging; CAD = computer-aided diagnosis; RF = random forest; SVM = support vector machine; MIL = multiple instance learning; ConvNet = convolutional network

From the Department of Oncology-Pathology, Karolinska Institutet, Stockholm, Sweden; Department of Clinical Pathology and Cytology, Karolinska University Laboratory, Stockholm, Sweden; School of Computer Science and Communication, KTH Royal Institute of Technology, Stockholm, Sweden; Science for Life Laboratory, Stockholm, Sweden; Stockholm South General Hospital, Stockholm, Sweden.

Submitted for Publication September 1, 2017; received submitted October 28, 2017; accepted for publication October 30, 2017.

Reprint requests: Johan Hartman, Department of Oncology-Pathology, Karolinska Institutet, CCK, SE-17176 Stockholm, Sweden; e-mail: johan.hartman@ki.se.

1931-5244/$ - see front matter

https://doi.org/10.1016/j.trsl.2017.10.010

BACKGROUND

Personalized cancer therapy—providing treatment tailored to the patient—will require sophisticated analyses on various levels, such as the genomic makeup of the tumor. In breast pathology, accurate biomarker assessment is vital for accurate therapy decisions. The complexity of and demand for accuracy in histopathologic breast cancer diagnosis is increasing. However, the lack of pathologists is an evident issue in most parts of the world. Computerized image analysis in histopathology of breast tumors holds promise to improve breast cancer diagnosis and help the pathologist reduce time-consuming tasks such as biomarker assessment. Several

biomarkers struggle with intra- and interobserver variability, which hampers its reproducibility. Digital image analysis (DIA) provides efficient tools to increase biomarker scoring reproducibility. Furthermore, development of computer-aided diagnosis and outcome prediction models using machine learning may facilitate clinical decision-making, in line with personalized cancer medicine.

Surgical pathology is a medical discipline completely dependent on microscopic images to diagnose diseases. With high capacity, whole-slide image scanners available, the digital workflow in surgical pathology is emerging, requiring advances in image analysis and providing opportunities for large collections of image data being used for machine learning algorithms. Groundbreaking improvements in computer-aided diagnostics and artificial intelligence (AI) will fundamentally change the way we diagnose diseases in the near future. The most promising advance in AI is machine learning, the science of making computers analyze and learn from data without human instruction. These technologies are commonly seen in areas such as spellcheck and development of self-driving cars, and are all carried out by neural network algorithms. Deep learning is a recent machine learning approach that uses biologically inspired networks to represent data through multiple levels of nonlinear modules that transform the previous representation into a higher, slightly more abstract representation. The compositional nature of the architecture allows deep neural networks to form highly complex and nonlinear representations that provide unprecedented discriminatory power. Deep networks have produced groundbreaking results in many important problems including image classification[1] and speech recognition.[2] Deep neural networks show decision-making capable of defeating world champion human Go players.[3] In medical imaging, deep neural network analysis of skin lesions has recently shown diagnostic accuracy on par with board-certified dermatologists.[4] In ophthalmologic assessment of retinal fundus images, neural networks recently showed high sensitivity and specificity for the detection of referable diabetic retinopathy.[5] An example of industry-driven innovation in this field is IBM building the AI platform Watson Health to aid radiologists interpret images and to produce even more sophisticated complex tasks to improve patient care.[6,7] The goal of AI applications is, however, not to replace radiologists or pathologists, but to make the diagnostic workflow more efficient and help evaluate and extract the most important information from the images, as well as to detect patterns not visible to the human eye.

In this review, we aim to summarize the recent developments in digital image analysis and in the use of AI in forms of machine learning in diagnostic breast pathology and to investigate why we need deep learning in histopathology.

THE PATHOLOGY OF BREAST CANCER

Breast cancer is a heterogeneous disease and the most common malignant disease in women worldwide, with nearly 1.7 million new cases in 2012.[8] More than 255,000 new cases are expected in 2017 in the United States alone.[9] The prognosis for breast cancer patients is highly variable, depending on several prognostic variables. Breast carcinomas comprise a wide range of morphologic phenotypes and are categorized into different histologic subtypes. Breast carcinoma arises from the mammary epithelium and causes a premalignant epithelial proliferation within the ducts, called ductal cancer in situ (DCIS). However, at some point the cancer cells may gain capacity to break through the basal membrane of the duct walls and infiltrate into surrounding tissues. When this happens, the disease is called invasive carcinoma.

Invasive carcinoma of no special type (NST) comprises the largest group of invasive breast cancers. Previously, invasive carcinoma NST was referred to as infiltrating ductal carcinoma or invasive ductal carcinoma not otherwise specified. In up to 80% of these tumors, foci of associated DCIS will be seen.[10,11] Invasive lobular carcinoma accounts for 5%–15% of invasive breast carcinoma, and appears with a more diffuse growth pattern of the cell infiltrate. Invasive carcinoma NST and lobular carcinoma together account for approximately 95% of all breast tumors. Apart from these, the World Health Organization has classified numerous other categories of carcinoma of the breast, which show characteristic morphology.[12,13]

Histopathology refers to the study of tissue specimen after the specimen has been fixed in formalin, paraffin embedded, and thin histologic tissue sections have been cut and mounted onto glass slides. The principal stain of tissue specimens is a combination of hematoxylin and eosin (H&E). H&E staining has been used for more than a century, and is still the standard for routine histopathologic diagnostics. In addition, specific techniques such as immunohistochemistry (IHC)[14-22] and in situ hybridization (ISH)[23] are often used to come to a complete diagnosis (Fig 1).

The morphologic assessment and tumor grading is manually performed by the pathologist through light microscope assessment of tissue sections. A tumor's histologic grade has shown to be a strong prognostic marker and is included in the pathologic reporting of every breast cancer specimen. The method for histologic grading by Bloom and Richardson has been modified most recently by Elston and Ellis.[24,25] The grading of invasive breast tumors is performed based on assessment of

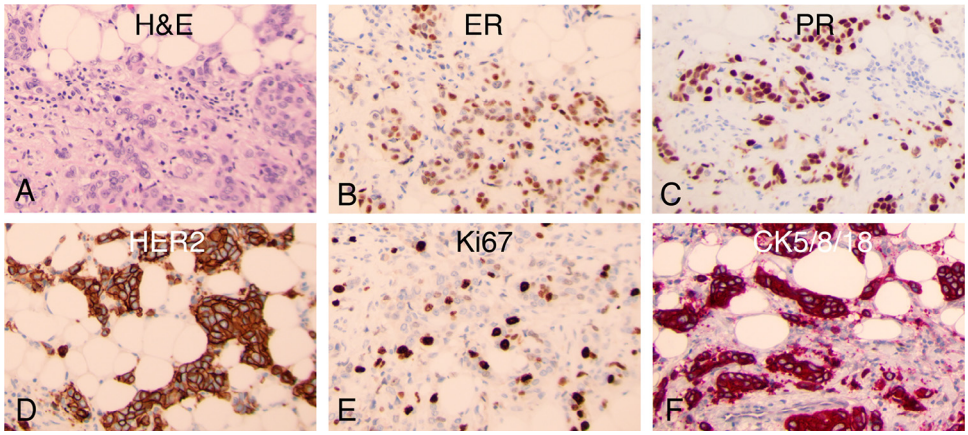

Fig 1. Breast carcinoma. (**A**) Hematoxylin and eosin tissue staining. Biomarkers immunohistochemically stained for estrogen receptor (ER; **B**), progesterone receptor (PR; **C**) HER2 **(D)**, and proliferation-associated protein Ki67 **(E).** (**F**) Immunohistochemical cytokeratin (CK) double staining for CK5 (brown) and CK8/CK18 (red) shows the absence of basal/myoepithelial cells (CK5) corresponding to invasive carcinoma.

3 morphologic features: tubular formation (glandular differentiation), nuclear pleomorphism, and mitotic counts, each given a value of 1–3. The combined scores give the overall tumor grade (I–III), currently named Nottingham histologic grade, because it originally was based on the Nottingham/Tenovus study 1973–1989.[25] This manual grading process is tedious and subjective, causing interobserver variations even among senior pathologists.[26]

The most important prognostic indicator in breast cancer is still its stage. Based on tumor size, regional lymph node status and distant metastasis, breast carcinomas are classified and staged according to the TNM system of malignant tumors.[13,27]

ASSESSMENT OF BIOMARKERS

The predictive and prognostic biomarkers estrogen receptor α (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2) and in some countries, the proliferation-associated nuclear protein Ki67 are routinely analyzed by IHC (Fig 1).[28-32] The American Society of Clinical Oncology/College of American Pathologists have published guideline recommendations for immunohistochemical testing of ER, PR, and HER2.[33,34] Approximately 80% of breast cancer patients have tumors that express ER, which is the only clinically used biomarker for endocrine therapy response.[35,36] In addition, co-expression of the ER-induced gene PR is considered a positive factor for response.[37] The hormone receptors, ER and PR, are assessed by manually counting the percentage of positive-stained tumor nuclei over the whole tumor section. Tumors with ≥1% or ≥10% stained nuclei, depending on guideline, are regarded as ER- or PR-positive.[34]

HER2/*neu* (c-*erb*B-2) is an oncogene that encodes a transmembrane glycoprotein with tyrosine kinase activity in the family of epidermal growth factor receptors. HER2 protein overexpression by HER2 gene amplification is seen in approximately 15% of early breast tumors and associated with poor prognosis.[38,39] But HER2 is also a target for anti-HER2 therapy that dramatically improves the outcome for this group of patients.[39] In routine clinical pathology, HER2 protein expression is first assessed by IHC analysis, and in those cases where the staining is equivocal, ISH analysis is performed to verify the HER2 gene amplification by a DNA probe.[40] The overexpression of HER2 is a good predictor of response to anti-HER2 therapy. However, diagnostic observer variability has been reported for HER2 scoring by pathologists.[41-45]

The proliferation-associated biomarker Ki67 is a nuclear protein expressed in all phases of the cell cycle, except for the G_0 (resting state). The Ki67 index is assessed by manual counting of the percentage of Ki67 stained nuclei in a "hot spot" region with the densest immunoreactivity out of at least 200 tumor cells. It can be used to distinguish between low- and high-proliferating tumors, which may be used to guide clinical decisions regarding adjuvant chemotherapy in ER-positive tumors.[46-48] There is robust evidence that Ki67 has clinical validity in breast cancer prognostication and therapy prediction.[46,49] Despite its importance in breast cancer, Ki67 index scoring has shown interlaboratory discordance and a lack of standardized assessment, which prohibits the recommendations for Ki67 assay in routine oncology.[28,46,50] Today, there is no consensus regarding which region to score or which cutoff to use for proliferation rates, and the IHC assessment of Ki67 assay

requires knowledge of local laboratory values.[28] As stated by Goldhirsch et al., "Analytic validity, clinical validity, and clinical utility are all required for optimal clinical application of tumor biomarkers."[29,51] Certain biomarkers are prone to intra- and interobserver variability, and this effect is especially evident for Ki67.[45,52,53] The St Gallen International Expert Consensus 2015 states that image analysis may help to reduce variability in Ki67 scoring.[28] The accuracy for the IHC-determined biomarkers is vital, because they together with the morphologic tumor characteristics (grade and stage) are used for guiding the therapeutic decisions.

Aside from standard clinicopathologic assessment, gene expression profiling can provide prognostic information.[54-57] High concordance and robustness of biomarker evaluations and gene expression profiling of breast tumors are of great importance in oncologic patient management. Identification of gene expression signatures and intrinsic subtypes based on global mRNA expression, as first described by Perou et al. in 2000, have provided promising alternatives to classify invasive breast tumors into high- and low-risk subtypes.[55,58] The intrinsic subtypes luminal A, luminal B, HER2-enriched, basal-like, and normal-like have been widely investigated by DNA microarray and hierarchical clustering analysis.[55,57-61] Multiparameter molecular marker assays, such as the Prosigna (Nanostring Technologies Inc, Seattle, WA),[62,63] BluePrint and MammaPrint (Agendia, Inc, Irvine, CA),[64,65] and Oncotype DX (Genomic Health Inc, San Francisco, CA)[66,67] can be used to define the subtypes and estimated risk of recurrence, but are unfortunately expensive and not available in less developed regions in the world. However, IHC measurements of ER, PR, Ki67, and ISH of HER2 can be used to approximate multigene testing, to obtain surrogate definition of subtypes.[68] The ER-expressing luminal A and B subtypes account for 70% of all breast tumors. The hormone receptor positive, low proliferating luminal A tumors generally have good prognosis and endocrine response,[69] whereas the high proliferating luminal B subtype show poorer prognosis and reduced sensitivity to endocrine therapy.[47] The HER2-enriched subgroup, which can be effectively targeted by anti-HER2-therapy, corresponds to ER-negative and HER2-positive tumors.[70] Triple-negative tumors correspond more or less to the basal-like subtype.[69]

WHY DIGITAL PATHOLOGY?

Around the world, pathology laboratories are faced with fixed health-care budgets in the face of growing numbers of patients from an aging population. At the same time, there is a shortage of trained pathologists, and patient demand for precision diagnostics and treatment is increasing. Efficient and cost-effective methods are highly desired to address these needs and to modernize routine pathology. The digitization of pathology data has opened the door to faster, more reproducible, and more precise diagnoses through computerized image analysis. Digital pathology is a dynamic, rapidly evolving, image-based discipline that incorporates acquisition, management, and interpretation of pathology information obtained from a digital scan of a glass slide. In several validation studies, the concordance between digital image diagnosis and conventional glass slide diagnosis has shown to be good to superior.[71-77] Digitalized images make consultation between expert pathologists easier, and facilitate automated image analysis. The goal of digital pathology is not to take over the pathologist's work, but to improve accuracy, reduce human error, and provide tools for a more efficient workflow and increased reproducibility. Other benefits of digitization include a reduction in the cost of handling glass slides and the ability to share data for education or long-distance consultations.[78] In addition, the image quality of a digital slide scan is preserved, whereas the staining in the physical sample fades over time.

Laboratories with integrated digital pathology workflows and interpretation are still sparse today. This stands in contrast to radiology, which in most countries underwent digitalization over the past decades. Picture archiving and communication systems have been in place for radiology departments for more than 20 years. Consequently, medical image analysis development has largely focused on digital radiology data, producing important advances with clinical implementations.[79,80] Technical limitations have slowed progress in digital histopathology—primarily because the images can be up to 10 times the size of radiology images. Histologic glass slides are converted to digital images using high-resolution whole-slide imaging (WSI) scanners. The entire process involves several steps: image acquisition, storage and management, annotation, and viewing or sharing. Several pathology laboratories are integrating WSI scanners into the routine workflow in an effort to digitalize the diagnostic workflow. However, challenges to find solutions that provide sufficient storage capacity with reasonable archiving costs remain. The digitalization of pathology data opens the door to quantitative computer-based image analysis, and could prove to be clinically important as a tool to accurately identify high-risk patients. Novel digital image analysis algorithms to improve therapy prediction and survival prognostication are of utmost value in the clinicopathologic setting. Current classification of biomarkers based on manual measurement is arbitrary: furthermore, manual measurement and counting of cell numbers hampers reproducibility. Depending on the downstream analyses that will be performed, manual annotation of the digitalized slides by a trained pathologist can be required, for example, to delineate cancer or

metastasis and exclude artifacts. That can be a time-consuming step. Computerized image analysis promises to improve reproducibility. However, there is controversy about how image analysis should be implemented.[28]

EXISTING DIGITAL IMAGE ANALYSIS METHODS

Pathology is an image-based discipline, traditionally with the light field microscope as the major working tool for image interpretation. Computerized image analysis software has been developed to aid the pathologist's evaluation of whole-slide images. Such software enables quantitative image analysis, in an effort to improve accuracy, reliability, reproducibility, and productivity. Computerized image analysis is not a novel research area in histopathology. Over the years, methods have been developed to reduce variations in image quality, for example, through color standardization, spatial filtering, denoising, or enhancement.[81] A significant amount of work has been devoted to the automatic segmentation of nuclei, for example, applying active contour models.[82] This is a general segmentation technique, which fits a deformable shape model to a given image.[83-85] Research has also concentrated on mitosis detection.[56] For a detailed review of object detection and segmentation, see Veta et al. and Gurcan et al.[86,87] In this review, we do not intend to cover all the previous techniques for image analysis in histopathology, but to give the current state of DIA focusing on breast pathology.

ImageJ, the free and accessible Java-based, user-friendly image analysis tool developed by the National Institutes of Health (Bethesda, MD), is arguably the most popular open tool for biomedical image analysis. It can readily be applied to quantify IHC or other specific tasks through different analysis and processing plugins. ImmunoRatio (University of Tampere, Finland) is a tool developed for automated analysis of IHC biomarker assessment (ER, PR, and Ki67). It provides a ratio of positively stained tumor nuclei area (Fig 2). It is available as a web application or as a plugin for ImageJ. ImmunoRatio has been used by various research groups, and studies computing the Ki67 labeling index have shown excellent concordance to manual assessment in breast cancer.[88,89] Similarly, the publicly available application ImmunoMembrane has been evaluated for HER2 IHC.[90]

Today, in industry, there is increasing competition for digital pathology image analysis solutions. This section gives a brief overview of the most common commercially available software for breast pathology. Roche VENTANA image analysis algorithm for IHC assays in breast pathology is a solution used to quantify breast panel biomarkers and to provide an integrated solution including antibody assays. In 2014, AstraZeneca acquired the imaging and data analysis technology company Definiens, and is incorporating their Tissue Phenomics software for clinical programs in immune-oncology and predictive biomarker discovery. Visiopharm (Hoersholm, Denmark) Virtual Double Staining techniques is using a pancytokeratin-stained tissue section, which is aligned to the IHC-stained biomarker of interest and enables automated detection of tumor regions (Fig 2).[91-93] The abovementioned commercial platforms operate on input from a WSI scanner, whereas the Aperio Digital Pathology (Leica Biosystems, Nussloch, Germany) platform operates by integrating a digital microscope with the image software. The Aperio Digital Pathology user-supervised platform has been compared with the automated Definiens Tissue Studio platform for classification of ER and PR IHC positivity.[94] TissueGnostics analysis software (Vienna, Austria) offers image analysis applications for clinical and research assessment of biomarkers in breast cancer. The rapid development of image analysis software and integrated solutions for histopathologic diagnostics will most certainly continue for the coming years, with close competition in the industry. If trends continue, these or similar software packages will become an integrated part of routine digitalized diagnostics.

Up to now, most research on digital image analysis has focused on quantifying biomarkers by IHC. DIA has shown excellent reproducibility, although limited to subsets with individual biomarkers or smaller cohorts.[93,95-97] Hartman et al. have recently shown the advantages in congruence to gene expression assays and the prognostic power of digital image analysis compared with current manual methods of biomarker assessments.[91] Automated image analysis has been applied to analyze several biomarkers, including HER2 expression, holding promise to reduce the need for additional ISH analysis in HER2 equivocal cases[95,98] and significantly reducing inter- and intraobserver variability.[41,42] Applying automated image analysis on IHC cytokeratin-stained sentinel node biopsies, negative (metastasis-free) samples can be eliminated with 100% sensitivity and used as a screening tool.[99] Furthermore, significant compression and scaling of large whole-slide images can be performed without comprising automated IHC biomarker assessment.[100]

An important next step will be to focus on the assessment of the basic H&E-stained tissue sections and the development of computer-aided diagnosis (CAD) algorithms for WSI.

MACHINE LEARNING IN DIGITAL PATHOLOGY IMAGE ANALYSIS

Recent breakthroughs in AI promise to fundamentally change the way we detect and treat breast cancer in

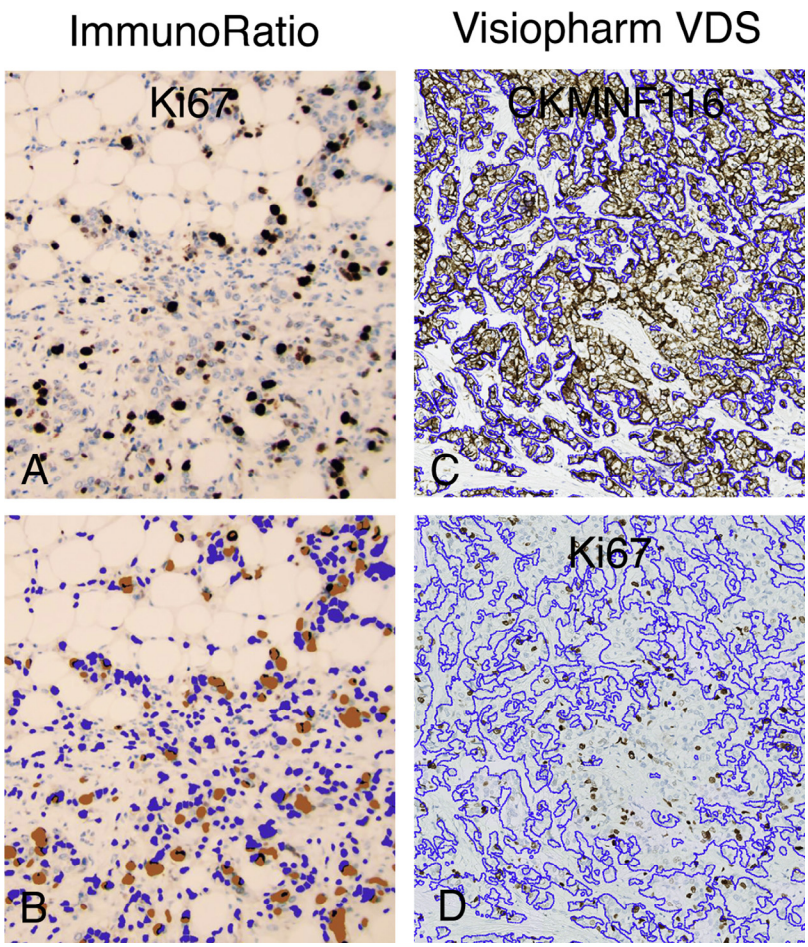

Fig 2. Digital image analysis for breast pathology. (**A-B**) Digital automated scoring of immunohistochemical (IHC) Ki67 using ImmunoRatio application for ImageJ. (**A**) Original image of IHC staining for Ki67 (brown nuclei). (**B**) Image showing staining components. Positive nuclei = orange. Negative nuclei = blue. (**C-D**) Visiopharm integrator software using virtual double staining (VDS) sandwich method for exclusion of non-epithelial cells and automated IHC Ki67 scoring. (**C**) IHC pancytokeratin CKMNF116 stained tumor cells automatically identified and outlined by blue line. (**D**) IHC Ki67 stained image aligned with the CKMNF116 image, used for scoring of Ki67 in tumor cells.

the near future. The difference between AI, machine learning, and deep learning is not always obvious to nonexperts. AI is an umbrella term encompassing the techniques for a machine to mimic or go beyond human intelligence, mainly in cognitive capabilities. AI includes a variety of subfields such as rule-based systems,[101] a classical approach to AI, in which the programmer explicitly encodes the knowledge provided by the task experts. In contrast, machine learning is a subfield of AI that applies statistical methods to learn to recognize patterns from a set of provided data without explicit human instruction. Deep learning is a recent machine learning approach that uses biologically inspired networks to represent data through multiple levels of simple but nonlinear modules that transform the previous representation into a higher, slightly more abstract representation. The compositional nature of the architecture allows deep neural networks to form highly complex and nonlinear representations that provide unprecedented discriminatory power. Deep networks have produced groundbreaking results in many important problems including image classification[1] and speech recognition.[2]

Computerized diagnostic systems in medicine[102] and technology in general,[103,104] have traditionally been rule-based. However, in the past years, we have witnessed pivotal progress in several aspects; the advent of powerful machine learning techniques, the advancement of graphics computational resources, and the ever-increasing digitization of medical data. These developments resulted in an explosion of interest in machine learning as these systems gradually replace classic image analysis techniques for automatic medical diagnosis.

The purpose of a machine learning algorithm is to use the provided task-related training data to learn a task,

which usually involves taking an input data and processing it in some way to produce a correct output. The training data contain different examples of the task's input. It can also include examples of the correct output, which is commonly referred to as labels. When the input data include the corresponding label for all the examples, it is called a supervised learning scenario. Supervised learning is currently the most common approach in digital histopathology.[105-122] The provided label for visual input data (eg, image) can correspond to either an entire image,[111-113,123-128] a window within the image,[108,115-117] or at the pixel level.[105-107,109,110,118-121] With the advent of deep learning methods that strongly benefit from pixel-level annotations, the latter is the most common type of problem being currently studied. Another popular scenario is when the training data come with a weaker form of annotation than what is expected as the output of the machine learning system. In work by Li et al.[113] and Zu et al.,[127] image-level labels are provided for histopathologic images of tumors or benign tissues; however, the algorithm is expected to produce pixel-level prediction of a cell being cancerous or not. This setup is referred to as weakly supervised learning.[113,115,127-129] Weakly supervised learning makes the annotation a simpler and less tedious task, but at the cost of the model becoming less accurate and robust. This trade-off of annotation costs and model accuracy is a recurring scenario in digital pathology because of the giga-pixel scale of the histopathologic image, perpetuating the prominence of weakly supervised learning. An even harder-to-learn setup is called unsupervised learning, where there is no example of the correct output provided with the training data. In unsupervised learning, the task is to look for patterns that we have minimal idea about a priori. For instance, Xu et al. use an unsupervised learning method to discover cancer subtypes by applying a clustering technique on histopathologic images of cancerous tissues.[127] Additionally, unsupervised learning is commonly employed for extracting the important measurements, or features, from the data.[107,130] There exist other variants of data supervision, which has less commonly been used in digital histopathology, such as reinforcement learning, semi-supervised learning, and self-supervised learning.

Before a machine learning algorithm can consume the training data to learn a model, it is required to make the data presentable to the algorithm. A representation method ideally discards the irrelevant information from the data and makes the potentially relevant information more accessible. In digital pathology, a typical histopathologic image can be $100{,}000 \times 100{,}000$ pixels. That is, each color image, as a computer program sees it, is 30 billion numbers. This enormity of the input data can easily cripple most of learning methods and warrants a representation technique that reduces this size. Until a few years ago, the most common types of image representations used in the literature as well as industry were hand-crafted by humans, such as scale-invariant feature transform,[131] speeded-up robust features,[132] and local binary patterns.[133] These hand-crafted representations are usually designed by domain experts who can hypothesize what type of image features can be important for their domain's tasks. In digital pathology, a combination of representations is usually used,[105,111,113,118,122,123,129,134] each capturing a different aspect of the image such as texture[105,111,113,123,134,135] and color.[113,134] Other works tried to learn the suitable representation instead of crafting them.[123,124,126,130,136-138] Most recently, with the resurgence of neural networks, representation learning has become the most prevalent method and achieved superior performance in various tasks.[106,107,111]

NON-DEEP LEARNING MACHINE LEARNING METHODS IN IMAGE ANALYSIS

Deep learning is currently the dominant technique for supervised learning from data. However, in the past decades, a diverse set of machine learning models has been proposed. Many of these methods have found their way into diagnostic medical imaging. A decision tree is used for supervised learning in classification and regression tasks.[139] It operates by breaking down the decision-making process into a series of consecutive tests. These tests form a tree-like structure, where each decision corresponds to a node. Starting from a single root node, each node applies a test on the input sample and based on the test outcome it forwards the sample to one of its children up the tree. When a sample reaches a leaf (a node with no children), the decision associated with that leaf is assigned to the sample. Decision trees are very efficient in training and inference, but are prone to overfitting—when the learned model becomes overly complex so that it fits the training data well but fails to generalize to unseen data.

A random forest (RF) uses an ensemble of decision trees to make its learning more robust to variations in data.[140,141] In RF, different random subsets of the training data are formed, and for each subset a decision tree is learned. An RF is very efficient thanks to the highly parallelizable nature of both its training and inference, and has been commonly used in digital pathology.[106,120,129]

Support vector machine (SVM) is a maximum-margin classification algorithm used for supervised learning.[142] It operates on training samples mapped to a representation space. For binary classification tasks (eg, tumor vs normal), it essentially means that it learns a linear separator in the representation space such that the training samples fall on either side of the line depending on their class. SVM further maximizes the distance of the closest

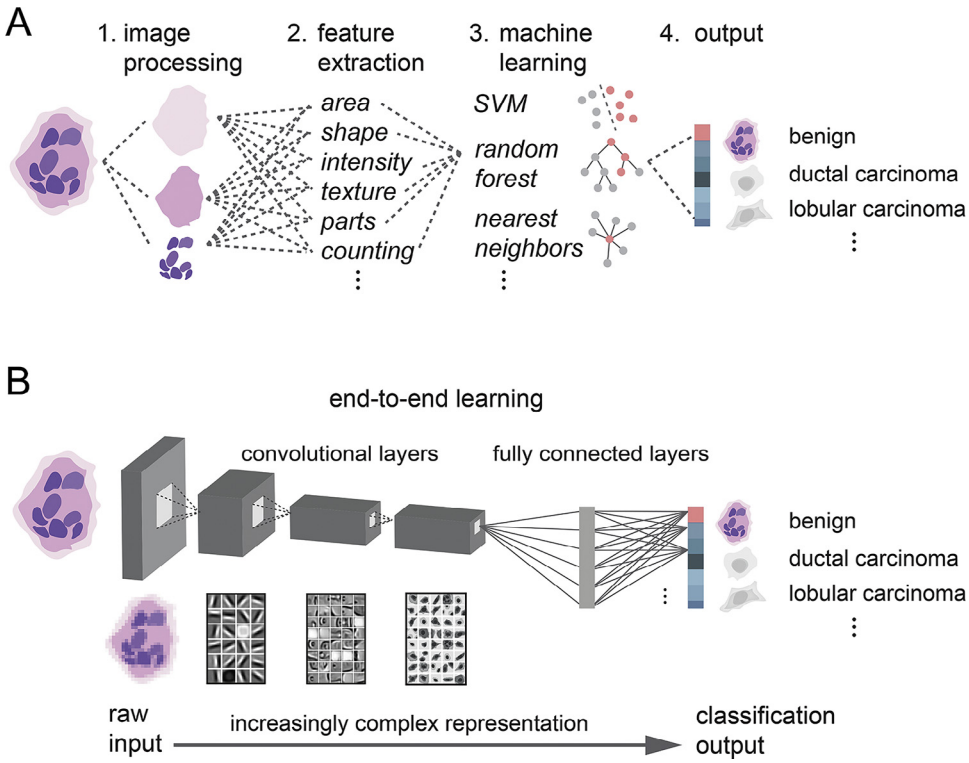

Fig 3. Deep learning vs traditional machine learning. (**A**) In the traditional paradigm, several steps requiring expert human knowledge are required to recognize cancer in images. First, image processing such as segmentation corrects the image and breaks it into manageable parts. Next, hand-crafted measurements, or features, are extracted from each part. A machine learning algorithm is provided those features as a vector, which it uses to learn a predictive model. (**B**) In contrast, deep learning is an end-to-end approach to learning that takes raw images as input and directly learns a model to produce the desired output. Deep learning uses biologically inspired networks to represent data through multiple levels of simple but nonlinear modules that transform the previous representation into a higher, slightly more abstract representation. The compositional nature of the architecture allows deep networks to form highly complex and nonlinear representations as each layer forms a more abstracted representation than the last. The result is a rich representation that provides unprecedented discriminatory power.

points (training samples) to the separation line. This ensures that there is a margin between samples of the 2 classes. Standard SVM can be naturally extended to multiclass scenarios.[143] SVMs are generally fast to train, and relatively efficient at inference. Various forms of SVM are commonly employed in image analysis applications of breast histopathology.[105,107,111,122-124,135]

Multiple instance learning (MIL) is a meta-algorithm dealing with a scenario where bags of samples are labeled instead of each sample individually. A positive bag means that at least one of the samples in the bag is positive and a negative bag means none of the samples in the bag is positive. MIL is usually applied to weakly supervised training data. In digital histopathology, it is mostly applied when we have labels for an image (bag of pixels) and not the pixels themselves, whereas we are interested in learning a model that can classify the pixels.[113,127-129] For instance, if the input is a dataset of images taken from tumor or benign specimens but the desired output is segmentation of the cancerous regions of an image, one can use MIL.

DEEP LEARNING

Deep learning is a method for layered end-to-end learning, and in contrast to other prevalent learning methods, deep learning requires minimal processing on the input data or the output values (Fig 3).[144] An end-to-end model takes in the raw data (eg, image pixels) and directly produces the desired output (eg, diagnosis), without an expert-designed feature extraction or representation step required by other learning methods. Deep learning involves modeling through multiple layers of nonlinear transformations. Each middle layer takes as input the output of the previous layer and transforms it by applying simple matrix operations using that layer's parameters. The first layer takes the raw data as input and the last layer produces the desired output (Fig 3). This design effectively unifies

the notions of representation learning and modeling such that it becomes indistinguishable at which layer the representation stops and the modeling starts. Although each layer's computations are simple, the power of deep learning comes from having multiple consecutive layers. It has been shown that deep models are more efficient than few layer models in approximating highly complex functions. In addition, the layer-wise transformation is compatible with our cognitive understanding of the compositional nature of an image—a scene consists of different objects, objects consist of parts, parts consist of motifs, and so forth. In that respect, earlier layers of a deep network model the more atomic constituents of an image, and as we go deeper more abstract concepts emerge.[145]

TYPES OF DEEP NETWORKS

There are several types of deep networks, with the most common ones being convolutional networks (ConvNet),[146] auto-encoders,[147] recurrent networks,[148] and adversarial networks.[149] ConvNet is composed of several convolutional layers; each of these layers applies the same local transformations at various locations of its input signal. A convolutional layer assumes that different concepts (eg, visual objects) can appear at any location in the input (eg. image). Convolutional layers, thus, have 2 properties, which significantly increase their efficiency: (1) the transformations are local and (2) the parameters are shared for different local regions. This profoundly increases the robustness of ConvNets to data variation. ConvNets are the most common types of deep learning used in digital breast pathology.[106,107,109-112,115,119,121,150,151]

Different architectures have been proposed for ConvNets. Some of these architectures have repeatedly been shown to be successful for image recognition tasks and became standards in the field. These architectures include inception,[108,150,152] residual networks,[109,153] visual geometry group,[106,150,151,154] multi-column networks,[155] densely connected networks,[156] and AlexNet.[1,112,150] Because of the unique size and nature of breast histopathology images, many works in this field design their own ConvNet architecture.[111,112,115,117,119,121]

Another type of deep network is an autoencoding network. Autoencoders are simple, usually shallow, architectures, which are used for learning a representation on a set of training data in an unsupervised fashion. This is usually done by encoding the data into a low-dimensional vector such that the important properties of the original data can be reconstructed from the low-dimensional vector. Autoencoders along with deep belief networks[157] and restricted Boltzmann machines[158] have been often used for representation learning on histopathologic images.[116,123,136-138]

Recently, a new architecture known as generative adversarial networks was proposed,[149] which has rapidly become popular because of its unprecedented ability to generate unseen examples.[159] Adversarial training involves 2 networks: a generator whose task is to generate realistic output, and a discriminator whose task is to distinguish between a synthetically generated output and a real example. The 2 networks compete—the generator's task is to perform so well at its task such that it completely fools the discriminator network. Adversarial training is starting to emerge in the field of digital pathology.[160,161]

COMMON PRACTICES IN DEEP LEARNING

A variety of different techniques have been established for training a deep learning model. Here we discuss a few key techniques commonly applied in digital pathology.

Deep learning is most successful when abundant labeled training data are available. ImageNet, a database commonly used to train networks for image classification, contains millions of images, and thousands of examples of every object type. Creating a dataset of this size is an enormous effort. Medical image datasets are usually much smaller because of patient privacy issues and the need for expert annotation. One way to circumvent this issue is to use transfer learning via pre-training the deep network on a completely different task[107,108,150] or a related task[106] and then fine-tuning the network's parameters on the small dataset at hand. Although this is a common setup for small medical datasets,[106-108,150] there are works who have managed to train a network successfully without pre-training.[109,119,121,150] Another approach to alleviate the lack of medical images is data augmentation using different image transformations that do not change its corresponding label. This is a common approach in digital histopathology using deep learning.[106,108-110,121,151] The augmentation can be done through flipping the images,[106,108-110,121,151] rotation,[106,108-110,121,151] cropping,[106,108,109,151] and less commonly via color and contrast distortion,[106,151] added blurriness,[110] scaling,[151] and elastic deformation.[110]

Another common technique includes processing the input image at different magnifications. Analyzing histopathologic images involves studying the phenotypic properties of individual cells as well as the growth patterns. This warrants a learning machinery, which processes the image at various scales. Thus, ensembles of multiple deep networks operating on different scales have been helpful in these applications.[108,110] In general, combining different deep networks for a single task is a common technique to increase the performance of the predictions.[106,108,121]

Finally, the trained deep network can be treated solely as a representation method and then be combined with other learning machineries.[106,107,111] Along that line, different machine learning techniques have been applied on top of a ConvNet's output, such as RF,[106] SVM,[107,111] voting,[108,109,117] probabilistic graphical models,[110] and additional adaptation layers.[107]

DEEP LEARNING FOR COMPUTATIONAL BREAST PATHOLOGY

Deep learning has generated a lot of excitement. The trend is to apply it to new domains, and pathology is not an exception. There is a high potential and demand for rapid advances in computational pathology. In breast pathology, apart from CAD, deep learning methods are being applied to complex pattern recognition tasks. Here, we intend to cover the leading research in computational breast pathology, with approaches for grading and detecting cancer as well as outcome prediction.

Deep learning has already been applied for detection of tubular formation,[162] nuclear pleomorphism,[162,163] and tumor grading.[162-164] Counting of mitotic figures, 1 of 3 parameters in grading of breast tumors, is a tedious and somewhat subjective assessment. Ciresan et al. was first to apply convolutional neural networks to counting of mitosis in breast cancer based on H&E sections.[121] Wang et al. proposed a ConvNets model combined with handcrafted features for mitosis detection[165] and later this method was combined with an RF classifier.[166] Another example combined segmentation-based features with neural networks and demonstrated precise detection of mitotic figures.[167] Several automated methods for mitosis detection have been proposed.[168,169]

Furthermore, deep learning algorithms such as ConvNets can facilitate classification of, for example, benign vs malignant breast tumors,[170] or detection of invasive carcinoma.[171] In a recent work by Araujo et al., a convolutional neural network could distinguish normal breast tissue, benign lesion, and in situ carcinoma from invasive carcinoma.[172] Automated methods have been proposed to distinguish intraductal lesions such as usual ductal hyperplasia from atypical ductal hyperplasia and DCIS, the latter 2 leading to surgical excision.[173] Cruz-Roa et al. proposed a multilevel neural network for automatic detection of invasive tumor extent in breast cancer.[174] Litjens et al. used deep learning model with fully ConvNets showing high performance in image classification in forms of detection of breast cancer metastasis in whole-slide images of sentinel lymph nodes.[175]

The Camelyon16 grand challenge was the first grand challenge on CAD in pathology using WSI. The data contained sentinel lymph node images of breast cancer patients, with the task being to detect metastasis. The best performing algorithms for this task all used deep learning, with a level of accuracy similar to that of a pathologist.[176-178]

Deep learning models have recently been applied to automated biomarker assessment in immunohistochemically stained breast tumor images. For example, they have been used for automatic scoring of IHC HER2[179] and automated ER scoring.[180] Furthermore, the stromal tissue surrounding the invasive tumor has not been as extensively studied, despite the vast cell types present in this tissue compartment and in direct relation to the tumor cells. However, deep learning for assessment of tumor-associated stroma and the diagnostic importance as a biomarker was recently shown.[106] In addition, deep learning methods have been used to quantify immune cell infiltration, so-called tumor infiltrating lymphocytes, in H&E-stained breast tumor images.[181]

Although deep learning has produced some promising results, before deep learning models are implemented in clinical decision-making, further studies are needed for validation and to assess their use.

Computer-aided prognosis is a promising field for machine learning algorithms, especially in cancer medicine, such as breast cancer and precision medicine. Computer-aided prognosis and prognostic models for breast cancer are based on histologic "features," molecular characteristics (gene expression profiling), and clinical data, and for example distinguish patients with more aggressive disease.[182,183] Machine learning algorithms have shown the ability to identify stromal morphologic structures that were not previously associated with prognostic outcome in breast cancer, and based on tissue microarray cohorts, an image-based model was developed to predict patient outcome.[184]

LIMITATIONS WITH DEEP LEARNING AND COMPUTER-AIDED DIAGNOSTICS

Because of its unprecedented performance in recognition tasks, deep learning has opened doors to technological advances in many fields including medical imaging. However, it still entails challenges, limitations, and concerns. This section discusses some of those that directly apply to the field of digital pathology.

Deep learning, in general, is perceived as a data-hungry learning method. The scale of the data available to most of the medical studies has so far been below the common standards in machine learning and computer vision. This can affect the performance of deep learning to medical data. So, the most common architectures and learning strategies in deep learning might not be immediately effective or optimal for medical image data. Moreover, standard deep architectures are computationally expensive and memory-heavy when training, and less

severely at deployment. The computational complexity gets aggravated for medical images because of their size (ranging from mega- to giga-pixels). These underline the importance of adapting the mainstream methods of deep learning toward the needs and settings of medical images.

Moreover, deep networks currently prevail in supervised learning, but this does not hold for weakly supervised or unsupervised learning scenarios. Supervised learning requires correctly labeled training data. In visual recognition, large datasets are created using crowdsourcing on common sense tasks (eg, is this an image of a cat?). Acquiring such data for medical questions can be problematic for several reasons. For histopathologic annotation, expert annotations from pathologists are required to do the job. Their time is a scarce and expensive resource. Another difficulty lies with the complexity of medical tasks; predicting a patient's outcome depends on various factors, and possibly the collected data may not give the complete picture—it may only be partially relevant to the outcome or even not relevant at all. Moreover, the most interesting estimation problems are the ones that the pathologists have internal disagreements about, which makes potentially noisy annotations. In addition, discovering novel biomarkers, a vital task for machine learning systems in medical image analysis, is an unsupervised learning problem, which deep learning is thus far not well suited for. Finally, because of the incredibly high resolution of some medical image data such as WSI, it can become prohibitively tedious to provide the necessary pixel-level annotations for localization tasks.

Another concern with deep learning, among other machine learning methods, is covariate shift, which refers to the scenario where the distribution of the input samples changes from that of the training time. This can become a vital reliability issue because all the performance metrics used to train and evaluate a machine learning system is drawn from the training distribution. If the training distribution no longer represents reality, there is no guarantee that the expected performance will still hold. In fact, it is possible to change the distribution of the input in such a way that a well-performing system completely fails at the new distribution. An important example in medical image analysis is when a model is trained only using images obtained from imaging equipment of a single vendor. The network then fails to respond correctly to images obtained from another vendor's equipment. One popular way that this phenomenon is studied in the context of deep learning is through the notion of adversarial examples. Most interestingly, it has been shown that image of a certain class can be minimally altered such that the network is completely fooled about its class, whereas the image's appearance remains the same to the human eye.[185] This concern is especially relevant to the field of medical imaging. Medical images can come from various equipment with different post-processing applied to them. Also, system reliability is a crucial aspect of medical applications.

Furthermore, a concern when analyzing histopathology images is color inconsistency and variations in tissue processing that may significantly affect the image analysis. To deal with these variations, caused by, for example, different staining techniques, staining procedures and manufacturers, and section thickness, different pre-processing methods are used for image color normalization.[81]

Finally, a common concern when acquiring medical data is the privacy of the patients and their medical records. Usually, that is resolved by giving restricted access to the individuals who are involved in an approved research project with an agreement not to publicize those data. However, releasing the trained model is usually not seen as a breach of this agreement. In that regard, some recent studies have shown that the data used to train a deep network might be partially retrievable by reverse engineering the network. However, techniques on how to use encrypted input data where reverse engineering become impossible have also been proposed.[186]

ADVANCES IN DIGITAL BREAST PATHOLOGY AND FUTURE PERSPECTIVES

In today's routine breast pathology, only a few pathology laboratories have adopted a digital workflow, which has limited the implementation of digital image analysis. However, there is evidence that automated image analysis increases reproducibility of biomarker assessment. It will not be long until digital pathology is the routine standard, as is the case in radiology. Deep learning algorithms for computer-aided prognosis may soon be able to retrieve similar prognostic data as gene expression profiling by identifying common patterns in different molecular subtypes of breast cancers. In this case, deep learning would be a cheaper and much faster alternative to gene expression profiling.

However, several challenges remain. Getting the full advantage of deep learning depends on having a full digital workflow, which is only slowly evolving because of the high costs and the dependence on solid IT support systems. The question of annotations remains: to what extent are annotations required and how can this be standardized? Might it be sufficient enough to separate DCIS from invasive tumor? Annotations limited to the training sets and to validate the networks in larger datasets.

Most studies have focused on invasive breast carcinoma, but little attention has been given to DCIS—could deep learning algorithms be trained to stratify DCIS lesions according to grade and aggressiveness and to

foresee which patient will progress to invasive cancer? A lot of work has been done on whole tumor sections, although core needle biopsies, with small amount of tissue on which the primary diagnosis often is given, can hold diagnostic challenges. Given a limited amount of breast tissue, DCIS can be difficult to distinguish from true invasive tumor cells. Appling machine learning on this aspect could aid the pathologist in this critical decision.

The main goal with the research in applied deep learning for breast pathology is to identify patterns not visible for the eye of a pathologist or so-called imaging biomarkers, new unknown biomarkers resulting from deep learning algorithms. Another promising avenue of research is to study patterns that correlate to molecular subtype, treatment response, and prognosis to refine the diagnostics in precision medicine. This is in line with the Precision Medicine Initiative launched by Barack Obama in 2015.[187] Stronger and robust biomarkers will likely not come from 1 level of data, but through combining genomic and histomorphologic information with neural networks that are able to provide robust prognostic and therapy-predictive information. Future studies will require well-established cohorts with multiomics data.

However, the challenge for machine learning is not to identify cancers or metastasis in images, because a trained pathologist can complete this task fast. What we really need deep learning for is to predict therapy response and prognostication of the tumors, and to combine with or complement genomics and transcriptomics for patient stratification. This is where the real challenge lies and it will be even more important to facilitate personalized cancer medicine. Collaborations between computer scientists, bioinformaticians, and research pathologists are of utmost importance for the development of relevant algorithms and computer-aided prognosis. In future research, we envision deep learning models combined with multiomics data for advanced precision medicine.

ACKNOWLEDGMENTS

Conflicts of Interest: All authors have read the journal's policy on disclosure of potential conflicts of interest and have none to declare. All authors have read the journal's authorship agreement and the manuscript has been reviewed and approved by all named authors. J.H. is member of the advisory board at Visiopharm A/S, Hoersholm, Denmark.

This work was supported with grants from the Swedish Society of Medicine, Swedish Society for Medical Research (SSMF), the Swedish Cancer Society, Stockholm Cancer Society, King Gustav V Jubilee Fund, Karolinska Institutet, Stockholm County Council Research Strategy Committee, and the Swedish Breast Cancer Association (BRO).

REFERENCES

1. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Adv Neural Inf Process Syst 2012:1097–105.
2. Hinton G, Deng L, Yu D, et al. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. IEEE Signal Process Mag 2012;29:82–97.
3. Silver D, Huang A, Maddison CJ, et al. Mastering the game of Go with deep neural networks and tree search. Nature 2016;529:484–9.
4. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017;542:115–8.
5. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 2016;316:2402–10.
6. Syeda-Mahmood T, Walach E, Beymer D, et al. Medical sieve: a cognitive assistant for radiologists and cardiologists. SPIE Medical Imaging; 2016: International Society for Optics and Photonics.
7. Anirudh R, Thiagarajan JJ, Bremer T, Kim H. Lung nodule detection using 3D convolutional neural networks trained on weakly labeled data. SPIE Medical Imaging; 2016: International Society for Optics and Photonics.
8. Stewart B, Wild C. World cancer report 2014. Lyon: International Agency for Research on Cancer; 2014.
9. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017. CA Cancer J Clin 2017;67:7–30.
10. Azzopardi AG. Problems in breast pathology. Classification of primary breast carcinoma. Philadelphia: W.B. Saunders, 1979.
11. Tavassoli FA. Pathology of the Breast. Stanford: Appleton and Lange, 1992.
12. Tavassoli FA, Devilee P. World Health Organization classification of tumours. In: Tavassoli FA, Devilee P, eds. Pathology and genetics of tumours of the breast and female genital organs. Lyon: IARC Press, 2003.
13. Lakhani SR. Lakhani SR, ed. WHO classification of tumours of the breast. Lyon: IARC Press, 2012.
14. Chu PG, Weiss LM. Keratin expression in human tissues and neoplasms. Histopathology 2002;40:403–39.
15. Perkins W, Campbell I, Leigh IM, MacKie RM. Keratin expression in normal skin and epidermal neoplasms demonstrated by a panel of monoclonal antibodies. J Cutan Pathol 1992;19:476–82.
16. Qureshi HS, Linden MD, Divine G, Raju UB. E-cadherin status in breast cancer correlates with histologic type but does not correlate with established prognostic parameters. Am J Clin Pathol 2006;125:377–85.
17. Berezowski K, Stastny JF, Kornstein MJ. Cytokeratins 7 and 20 and carcinoembryonic antigen in ovarian and colonic carcinoma. Mod Pathol 1996;9:426–9.
18. Lagendijk JH, Mullink H, Van Diest PJ, Meijer GA, Meijer CJ. Tracing the origin of adenocarcinomas with unknown primary using immunohistochemistry: differential diagnosis between colonic and ovarian carcinomas as primary sites. Hum Pathol 1998;29:491–7.
19. Loy TS, Calaluce RD. Utility of cytokeratin immunostaining in separating pulmonary adenocarcinomas from colonic adenocarcinomas. Am J Clin Pathol 1994;102:764–7.
20. Nadji M, Tabei SZ, Castro A, et al. Prostatic-specific antigen: an immunohistologic marker for prostatic neoplasms. Cancer 1981;48:1229–32.
21. Gould VE. Synaptophysin. A new and promising pan-neuroendocrine marker. Arch Pathol Lab Med 1987;111:791–4.

22. Cao Y, Schlag PM, Karsten U. Immunodetection of epithelial mucin (MUC1, MUC3) and mucin-associated glycotopes (TF, Tn, and sialosyl-Tn) in benign and malignant lesions of colonic epithelium: apolar localization corresponds to malignant transformation. Virchows Arch 1997;431:159–66.
23. Naber SP. Molecular pathology—diagnosis of infectious disease. N Engl J Med 1994;331:1212–5.
24. Scarff RW, Torloni H. Histological typing of breast tumours (International Histological Classification of Tumours, no. 2). Geneva: World Health Organization; 1968.
25. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology 1991;19:403–10.
26. Meyer JS, Alvarez C, Milikowski C, et al. Breast carcinoma malignancy grading by Bloom-Richardson system vs proliferation index: reproducibility of grade and advantages of proliferation index. Mod Pathol 2005;18:1067–78.
27. Brierley JD, Gospodarowicz MK, Wittekind C. Brierley JD, Gospodarowicz MK, Wittekind C, eds. TNM classification of malignant tumours. 8th ed. Oxford, UK: John Wiley & Sons, Ltd, 2017.
28. Coates AS, Winer EP, Goldhirsch A, et al. Tailoring therapies—improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015. Ann Oncol 2015;26:1533–46.
29. Goldhirsch A, Winer EP, Coates AS, et al. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. Ann Oncol 2013;24:2206–23.
30. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thurlimann B, Senn HJ. Strategies for subtypes—dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. Ann Oncol 2011;22:1736–47.
31. Harris L, Fritsche H, Mennel R, et al. American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. J Clin Oncol 2007;25:5287–312.
32. Van Poznak C, Somerfield MR, Bast RC, et al. Use of biomarkers to guide decisions on systemic therapy for women with metastatic breast cancer: American Society of Clinical Oncology Clinical Practice Guideline. J Clin Oncol 2015;33:2695–704.
33. Harris LN, Ismaila N, McShane LM, et al. Use of biomarkers to guide decisions on adjuvant systemic therapy for women with early-stage invasive breast cancer: American Society of Clinical Oncology Clinical Practice Guideline. J Clin Oncol 2016;34:1134–50.
34. Hammond ME, Hayes DF, Dowsett M, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. Arch Pathol Lab Med 2010;134:907–22.
35. Hawkins RA, Roberts MM, Forrest AP. Oestrogen receptors and breast cancer: current status. Br J Surg 1980;67:153–69.
36. Barnes DM, Hanby AM. Oestrogen and progesterone receptors in breast cancer: past, present and future. Histopathology 2001;38:271–4.
37. Li CI, Daling JR, Malone KE. Incidence of invasive breast cancer by hormone receptor status from 1992 to 1998. J Clin Oncol 2003;21:28–34.
38. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science 1987;235:177–82.
39. Wolff AC, Hammond ME, Hicks DG, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. J Clin Oncol 2013;31:3997–4013.
40. Grabau D, Quality- and Standardization Committee (KVAST). Quality- and standadization document for breast tumors. Alingsås, Sweden: Swedish Pathological Society/Swedish Society of Clinical Cytology; 2014.
41. Gavrielides MA, Gallas BD, Lenz P, Badano A, Hewitt SM. Observer variability in the interpretation of HER2/neu immunohistochemical expression with unaided and computer-aided digital microscopy. Arch Pathol Lab Med 2011;135:233–42.
42. Bloom K, Harrington D. Enhanced accuracy and reliability of HER-2/neu immunohistochemical scoring using digital microscopy. Am J Clin Pathol 2004;121:620–30.
43. Roche PC, Suman VJ, Jenkins RB, et al. Concordance between local and central laboratory HER2 testing in the breast intergroup trial N9831. J Natl Cancer Inst 2002;94:855–7.
44. Perez EA, Suman VJ, Davidson NE, et al. HER2 testing by local, central, and reference laboratories in specimens from the North Central Cancer Treatment Group N9831 intergroup adjuvant trial. J Clin Oncol 2006;24:3032–8.
45. Bueno-de-Mesquita JM, Nuyten DS, Wesseling J, van Tinteren H, Linn SC, van de Vijver MJ. The impact of inter-observer variation in pathological assessment of node-negative breast cancer on clinical risk assessment and patient selection for adjuvant systemic treatment. Ann Oncol 2010;21:40–7.
46. Dowsett M, Nielsen TO, A'Hern R, et al. Assessment of Ki67 in breast cancer: recommendations from the International Ki67 in Breast Cancer working group. J Natl Cancer Inst 2011;103:1656–64.
47. Criscitiello C, Disalvatore D, De Laurentiis M, et al. High Ki-67 score is indicative of a greater benefit from adjuvant chemotherapy when added to endocrine therapy in luminal B HER2 negative and node-positive breast cancer. Breast 2014;23:69–75.
48. Curigliano G, Burstein HJ, P Winer E, 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;28:1700–12.
49. de Azambuja E, Cardoso F, de Castro G Jr, et al. Ki-67 as prognostic marker in early breast cancer: a meta-analysis of published studies involving 12,155 patients. Br J Cancer 2007;96:1504–13.
50. Polley MY, Leung SC, Gao D, et al. An international study to increase concordance in Ki67 scoring. Mod Pathol 2015;28:778–86.
51. Hayes DF. From genome to bedside: are we lost in translation? Breast 2013;22(Suppl 2):S22–6.
52. Polley MY, Leung SC, McShane LM, et al. An international Ki67 reproducibility study. J Natl Cancer Inst 2013;105:1897–906.
53. Varga Z, Diebold J, Dommann-Scherrer C, et al. How reliable is Ki-67 immunohistochemistry in grade 2 breast carcinomas? A QA study of the Swiss Working Group of Breast- and Gynecopathologists. PLoS ONE 2012;7:e37379.
54. van 't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002;415:530–6.
55. Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci USA 2001;98:10869–74.

56. Loi S, Haibe-Kains B, Desmedt C, et al. Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade. J Clin Oncol 2007;25:1239–46.
57. Parker JS, Mullins M, Cheang MC, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol 2009;27:1160–7.
58. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. Nature 2000;406:747–52.
59. Fan C, Oh DS, Wessels L, et al. Concordance among gene-expression-based predictors for breast cancer. N Engl J Med 2006;355:560–9.
60. Hu Z, Fan C, Oh DS, et al. The molecular portraits of breast tumors are conserved across microarray platforms. BMC Genomics 2006;7:96.
61. Sorlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc Natl Acad Sci USA 2003;100:8418–23.
62. Bastien RR, Rodriguez-Lescure A, Ebbert MT, et al. PAM50 breast cancer subtyping by RT-qPCR and concordance with standard clinical molecular markers. BMC Med Genomics 2012;5:44.
63. Nielsen TO, Parker JS, Leung S, et al. A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor-positive breast cancer. Clin Cancer Res 2010;16:5222–32.
64. Bayraktar S, Royce M, Stork-Sloots L, de Snoo F, Gluck S. Molecular subtyping predicts pathologic tumor response in early-stage breast cancer treated with neoadjuvant docetaxel plus capecitabine with or without trastuzumab chemotherapy. Med Oncol 2014;31:163.
65. Nguyen B, Cusumano PG, Deck K, et al. Comparison of molecular subtyping with BluePrint, MammaPrint, and TargetPrint to local clinical subtyping in breast cancer patients. Ann Surg Oncol 2012;19:3257–63.
66. Dowsett M, Sestak I, Lopez-Knowles E, et al. Comparison of PAM50 risk of recurrence score with oncotype DX and IHC4 for predicting risk of distant recurrence after endocrine therapy. J Clin Oncol 2013;31:2783–90.
67. Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 2004;351:2817–26.
68. Cheang MC, Chia SK, Voduc D, et al. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. J Natl Cancer Inst 2009;101:736–50.
69. Guiu S, Michiels S, Andre F, et al. Molecular subclasses of breast cancer: how do we define them? The IMPAKT 2012 Working Group Statement. Ann Oncol 2012;23:2997–3006.
70. de Azambuja E, Holmes AP, Piccart-Gebhart M, et al. Lapatinib with trastuzumab for HER2-positive early breast cancer (NeoALTTO): survival outcomes of a randomised, open-label, multicentre, phase 3 trial and their association with pathological complete response. Lancet Oncol 2014;15:1137–46.
71. Ho J, Parwani AV, Jukic DM, Yagi Y, Anthony L, Gilbertson JR. Use of whole slide imaging in surgical pathology quality assurance: design and pilot validation studies. Hum Pathol 2006;37:322–31.
72. Helin H, Lundin M, Lundin J, et al. Web-based virtual microscopy in teaching and standardizing Gleason grading. Hum Pathol 2005;36:381–6.
73. Fine JL, Grzybicki DM, Silowash R, et al. Evaluation of whole slide image immunohistochemistry interpretation in challenging prostate needle biopsies. Hum Pathol 2008;39:564–72.
74. Lopez AM, Graham AR, Barker GP, et al. Virtual slide telepathology enables an innovative telehealth rapid breast care clinic. Hum Pathol 2009;40:1082–91.
75. Velez N, Jukic D, Ho J. Evaluation of 2 whole-slide imaging applications in dermatopathology. Hum Pathol 2008;39:1341–9.
76. Evans AJ, Chetty R, Clarke BA, et al. Primary frozen section diagnosis by robotic microscopy and virtual slide telepathology: the University Health Network experience. Hum Pathol 2009;40:1070–81.
77. Wienert S, Beil M, Saeger K, Hufnagl P, Schrader T. Integration and acceleration of virtual microscopy as the key to successful implementation into the routine diagnostic process. Diagn Pathol 2009;4:3.
78. Jara-Lazaro AR, Thamboo TP, Teh M, Tan PH. Digital pathology: exploring its applications in diagnostic surgical pathology practice. Pathology 2010;42:512–8.
79. Warren Burhenne LJ, Wood SA, D'Orsi CJ, et al. Potential contribution of computer-aided detection to the sensitivity of screening mammography. Radiology 2000;215:554–62.
80. Shiraishi J, Li Q, Appelbaum D, Doi K. Computer-aided diagnosis and artificial intelligence in clinical imaging. Semin Nucl Med 2011;41:449–62.
81. Khan AM, Rajpoot N, Treanor D, Magee D. A nonlinear mapping approach to stain normalization in digital histopathology images using image-specific color deconvolution. IEEE Trans Biomed Eng 2014;61:1729–38.
82. Kass M, Witkin A, Terzopoulos D. Snakes: active contour models. Int J Comput Vis 1988;1:321–31.
83. Jing J, Wan T, Cao J, Qin Z. An improved hybrid active contour model for nuclear segmentation on breast cancer histopathology. Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on; 2016: IEEE.
84. Ali S, Madabhushi A. Active contour for overlap resolution using watershed based initialization (ACOReW): Applications to histopathology. Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on; 2011: IEEE.
85. Ali S, Madabhushi A. An integrated region-, boundary-, shape-based active contour for multiple object overlap resolution in histological imagery. IEEE Trans Med Imaging 2012;31:1448–60.
86. Veta M, Pluim JP, van Diest PJ, Viergever MA. Breast cancer histopathology image analysis: a review. IEEE Trans Biomed Eng 2014;61:1400–11.
87. Gurcan MN, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, Yener B. Histopathological image analysis: a review. IEEE Rev Biomed Eng 2009;2:147–71.
88. Tuominen VJ, Ruotoistenmaki S, Viitanen A, Jumppanen M, Isola J. ImmunoRatio: a publicly available web application for quantitative image analysis of estrogen receptor (ER), progesterone receptor (PR), and Ki-67. Breast Cancer Res 2010;12:12.
89. Yeo MK, Kim HE, Kim SH, Chae BJ, Song BJ, Lee A. Clinical usefulness of the free web-based image analysis application ImmunoRatio for assessment of Ki-67 labelling index in breast cancer. J Clin Pathol 2017;70:715–9.
90. Tuominen VJ, Tolonen TT, Isola J. ImmunoMembrane: a publicly available web application for digital image analysis of HER2 immunohistochemistry. Histopathology 2012;60:758–67.
91. Stalhammar G, Fuentes Martinez N, Lippert M, et al. Digital image analysis outperforms manual biomarker assessment in breast cancer. Mod Pathol 2016;29:318–29.
92. Lykkegaard Andersen N, Brugmann A, Lelkaitis G, Nielsen S, Friis Lippert M, Vyberg M. Virtual Double Staining: a digital approach to immunohistochemical quantification of estrogen

receptor protein in breast carcinoma specimens. Appl Immunohistochem Mol Morphol 2017;doi:10.1097/PAI.0000000000000502.
93. Roge R, Riber-Hansen R, Nielsen S, Vyberg M. Proliferation assessment in breast carcinomas using digital image analysis based on virtual Ki67/cytokeratin double staining. Breast Cancer Res Treat 2016;158:11–9.
94. Ahern TP, Beck AH, Rosner BA, et al. Continuous measurement of breast tumour hormone receptor expression: a comparison of two computational pathology platforms. J Clin Pathol 2017;70:428–34.
95. Holten-Rossing H, Moller Talman ML, Kristensson M, Vainer B. Optimizing HER2 assessment in breast cancer: application of automated image analysis. Breast Cancer Res Treat 2015;152:367–75.
96. Rizzardi AE, Johnson AT, Vogel RI, et al. Quantitative comparison of immunohistochemical staining measured by digital image analysis versus pathologist visual scoring. Diagn Pathol 2012;7:42.
97. Nielsen PS, Riber-Hansen R, Jensen TO, Schmidt H, Steiniche T. Proliferation indices of phosphohistone H3 and Ki67: strong prognostic markers in a consecutive cohort with stage I/II melanoma. Mod Pathol 2013;26:404–13.
98. Brugmann A, Eld M, Lelkaitis G, et al. Digital image analysis of membrane connectivity is a robust measure of HER2 immunostains. Breast Cancer Res Treat 2012;132:41–9.
99. Holten-Rossing H, Talman MM, Jylling AMB, Laenkholm AV, Kristensson M, Vainer B. Application of automated image analysis reduces the workload of manual screening of sentinel Lymph node biopsies in breast cancer. Histopathology 2017;doi:10.1111/his.13305.
100. Konsti J, Lundin M, Linder N, et al. Effect of image compression and scaling on automated scoring of immunohistochemical stainings and segmentation of tumor epithelium. Diagn Pathol 2012;7:29.
101. Buchanan BG, Shortliffe EH. Rule-based expert systems: Addison-Wesley Reading, MA; 1984.
102. Olson WH, Kaemmerer WF. Prioritized rule based method and apparatus for diagnosis and treatment of arrhythmias. Google Patents; 1996.
103. Kramer MA, Palowitch B. A rule-based approach to fault diagnosis using the signed directed graph. AIChE J 1987;33:1067–78.
104. Xu J. Rule-based automatic software performance diagnosis and improvement. Perform Eval 2012;69:525–50.
105. Azimi V, Chang YH, Thibault G, et al. Breast cancer histopathology image analysis pipeline for tumor purity estimation. Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on; 2017: IEEE.
106. Ehteshami Bejnordi B, Linz J, Glass B, et al. Deep learning-based assessment of tumor-associated stroma for diagnosing breast cancer in histopathology images. arXiv preprint arXiv:170205803. 2017.
107. Song Y, Zou JJ, Chang H, Cai W. Adapting fisher vectors for histopathology image classification. Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on; 2017: IEEE.
108. Das K, Karri SPK, Roy AG, Chatterjee J, Sheet D. Classifying histopathology whole-slides using fusion of decisions from deep convolutional network on a collection of random multi-views at multi-magnification. Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on; 2017: IEEE.
109. Wollmann T, Rohr K. Deep residual Hough voting for mitotic cell detection in histopathology images. Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on; 2017: IEEE.
110. Naylor P, Laé M, Reyal F, Walter T. Nuclei segmentation in histopathology images using deep neural networks. Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on; 2017: IEEE.
111. Cao J, Qin Z, Jing J, Chen J, Wan T. An automatic breast cancer grading method in histopathological images based on pixel-, object-, and semantic-level features. Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on; 2016: IEEE.
112. Spanhol FA, Oliveira LS, Petitjean C, Heutte L. Breast cancer histopathological image classification using convolutional neural networks. Neural Networks (IJCNN), 2016 International Joint Conference on; 2016: IEEE.
113. Li W, Zhang J, McKenna SJ. Multiple instance cancer detection by boosting regularised trees. International Conference on Medical Image Computing and Computer-Assisted Intervention; 2015: Springer.
114. Paul A, Dey A, Mukherjee DP, Sivaswamy J, Tourani V. Regenerative random forest with automatic feature selection to detect mitosis in histopathological breast cancer images. International Conference on Medical Image Computing and Computer-Assisted Intervention; 2015: Springer.
115. Xie Y, Xing F, Kong X, Su H, Yang L. Beyond classification: structured regression for robust cell detection using convolutional neural network. International Conference on Medical Image Computing and Computer-Assisted Intervention; 2015: Springer.
116. Su H, Xing F, Kong X, Xie Y, Zhang S, Yang L. Robust cell detection and segmentation in histopathological images using sparse reconstruction and stacked denoising autoencoders. International Conference on Medical Image Computing and Computer-Assisted Intervention; 2015: Springer.
117. Xie Y, Kong X, Xing F, Liu F, Su H, Yang L. Deep voting: A robust approach toward nucleus localization in microscopy images. International Conference on Medical Image Computing and Computer-Assisted Intervention; 2015: Springer.
118. Akbar S, Jordan L, Thompson AM, McKenna SJ. Tumor localization in tissue microarrays using rotation invariant superpixel pyramids. Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on; 2015: IEEE.
119. Su H, Liu F, Xie Y, Xing F, Meyyappan S, Yang L. Region segmentation in histopathological breast cancer images using deep convolutional neural network. Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on; 2015: IEEE.
120. Peter L, Mateus D, Chatelain P, et al. Leveraging random forests for interactive exploration of large histological images. International Conference on Medical Image Computing and Computer-Assisted Intervention; 2014: Springer.
121. Ciresan DC, Giusti A, Gambardella LM, Schmidhuber J. Mitosis detection in breast cancer histology images with deep neural networks. Med Image Comput Comput Assist Interv 2013;16:411–8.
122. Arteta C, Lempitsky V, Noble JA, Zisserman A. Learning to detect cells using non-overlapping extremal regions. International Conference on Medical Image Computing and Computer-Assisted Intervention; 2012: Springer.
123. Song Y, Li Q, Huang H, Feng D, Chen M, Cai W. Histopathology Image Categorization with Discriminative Dimension Reduction of Fisher Vectors. European Conference on Computer Vision; 2016: Springer.
124. Vu TH, Mousavi HS, Monga V, Rao UA, Rao G. DFDL: Discriminative feature-oriented dictionary learning for histopathological image classification. Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on; 2015: IEEE.
125. Khan AM, Sirinukunwattana K, Rajpoot N. Geodesic geometric mean of regional covariance descriptors as an image-level

descriptor for nuclear atypia grading in breast histology images. International Workshop on Machine Learning in Medical Imaging; 2014: Springer.
126. Srinivas U, Mousavi H, Jeon C, Monga V, Hattel A, Jayarao B. SHIRC: A simultaneous sparsity model for histopathological image representation and classification. Biomedical Imaging (ISBI), 2013 IEEE 10th International Symposium on; 2013: IEEE.
127. Xu Y, Zhu J-Y, Chang E, Tu Z. Multiple clustered instance learning for histopathology cancer image classification, segmentation and clustering. Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on; 2012: IEEE.
128. Xu Y, Zhang J, Eric I, Chang C, Lai M, Tu Z. Context-constrained multiple instance learning for histopathology image segmentation. International Conference on Medical Image Computing and Computer-Assisted Intervention; 2012: Springer.
129. Cheplygina V, Sørensen L, Tax DM, de Bruijne M, Loog M. Label stability in multiple instance learning. International Conference on Medical Image Computing and Computer-Assisted Intervention; 2015: Springer.
130. Chang H, Zhou Y, Spellman P, Parvin B. Stacked predictive sparse coding for classification of distinct regions in tumor histopathology. Proceedings of the IEEE International Conference on Computer Vision; 2013.
131. Lowe DG. Object recognition from local scale-invariant features. Computer vision, 1999 The proceedings of the seventh IEEE international conference on; 1999: Ieee.
132. Bay H, Tuytelaars T, Van Gool L. Surf: Speeded up robust features. Computer vision–ECCV 2006. 2006:404–17.
133. Ojala T, Pietikainen M, Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 2002;24: 971–87.
134. Cheikh BB, Elie N, Plancoulaine B, Bor-Angelier C, Racoceanu D. Spatial interaction analysis with graph based mathematical morphology for histopathology. Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on; 2017: IEEE.
135. Doyle S, Agner S, Madabhushi A, Feldman M, Tomaszewski J. Automated grading of breast cancer histopathology using spectral clustering with textural and architectural image features. Biomedical Imaging: From Nano to Macro, 2008 ISBI 2008 5th IEEE International Symposium on; 2008: IEEE.
136. Xu J, Xiang L, Liu Q, et al. Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images. IEEE Trans Med Imaging 2016;35:119–30.
137. Cruz-Roa AA, Ovalle JEA, Madabhushi A, Osorio FAG. A deep learning architecture for image representation, visual interpretability and automated basal-cell carcinoma cancer detection. International Conference on Medical Image Computing and Computer-Assisted Intervention; 2013: Springer.
138. Nayak N, Chang H, Borowsky A, Spellman P, Parvin B. Classification of tumor histopathology via sparse feature learning. Biomedical Imaging (ISBI), 2013 IEEE 10th International Symposium on; 2013: IEEE.
139. Quinlan JR. Induction of decision trees. Machine learning 1986;1:81–106.
140. Breiman L. Random forests. Machine learning 2001;45:5–32.
141. Liaw A, Wiener M. Classification and regression by randomForest. R News 2002;2:18–22.
142. Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995;20:273–97.
143. Weston J, Watkins C. Multi-class support vector machines. Technical Report CSD-TR-98-04, Department of Computer Science, Royal Holloway, University of London, May; 1998.
144. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521: 436–44.
145. Bengio Y, Courville A, Vincent P. Representation learning: a review and new perspectives. IEEE Trans Pattern Anal Mach Intell 2013;35:1798–828.
146. LeCun Y, Kavukcuoglu K, Farabet C. Convolutional networks and applications in vision. Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on; 2010: IEEE.
147. Kingma DP, Welling M. Auto-encoding variational Bayes. arXiv preprint arXiv:13126114. 2013.
148. Graves A. Supervised sequence labelling with recurrent neural networks. Springer, 2012.
149. Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets. Adv Neural Inf Process Syst; 2014.
150. Bayramoglu N, Heikkilä J. Transfer learning for cell nuclei classification in histopathology images. Computer Vision–ECCV 2016 Workshops; 2016: Springer.
151. Veta M, Van Diest PJ, Pluim JP. Cutting out the middleman: measuring nuclear area in histopathology slides without segmentation. International Conference on Medical Image Computing and Computer-Assisted Intervention; 2016: Springer.
152. Szegedy C, Ioffe S, Vanhoucke V, Alemi AA. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. AAAI; 2017.
153. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition; 2016.
154. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:14091556. 2014.
155. Ciregan D, Meier U, Schmidhuber J. Multi-column deep neural networks for image classification. Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on; 2012: IEEE.
156. Huang G, Liu Z, Weinberger KQ, van der Maaten L. Densely connected convolutional networks. arXiv preprint arXiv: 160806993. 2016.
157. Hinton GE. Deep belief networks. Scholarpedia 2009;4:5947.
158. Salakhutdinov R, Hinton G. Deep boltzmann machines. Artificial Intelligence and Statistics; 2009.
159. Huang R, Zhang S, Li T, He R. Beyond Face Rotation: Global and Local Perception GAN for Photorealistic and Identity Preserving Frontal View Synthesis. arXiv preprint arXiv: 170404086. 2017.
160. Lafarge MW, Pluim JP, Eppenhof KA, Moeskops P, Veta M. Domain-adversarial neural networks to address the appearance variability of histopathology images. arXiv preprint arXiv: 170706183. 2017.
161. Udrea A, Mitra GD. Generative Adversarial Neural Networks for Pigmented and Non-Pigmented Skin Lesions Detection in Clinical Images. Control Systems and Computer Science (CSCS), 2017 21st International Conference on; 2017: IEEE.
162. Dalle JR, Leow WK, Racoceanu D, Tutac AE, Putti TC. Automatic breast cancer grading of histopathological images. Conf Proc IEEE Eng Med Biol Soc 2008;2008:3052–5.
163. Petushi S, Garcia FU, Haber MM, Katsinis C, Tozeren A. Large-scale computations on histology images reveal grade-differentiating parameters for breast cancer. BMC Med Imaging 2006;6:14.
164. Basavanhally A, Ganesan S, Feldman M, et al. Multi-field-of-view framework for distinguishing tumor grade in ER+ breast cancer from entire histopathology slides. IEEE Trans Biomed Eng 2013;60:2089–99.

165. Wang H, Cruz-Roa A, Basavanhally A, et al. Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features. J Med Imaging (Bellingham) 2014;1:034003.
166. Veta M, van Diest PJ, Willems SM, et al. Assessment of algorithms for mitosis detection in breast cancer histopathology images. Med Image Anal 2015;20:237–48.
167. Malon CD, Cosatto E. Classification of mitotic figures with convolutional neural networks and seeded blob features. J Pathol Inform 2013;4:9.
168. Irshad H. Automated mitosis detection in histopathology using morphological and multi-channel statistics features. J Pathol Inform 2013;4:10.
169. Roux L, Racoceanu D, Lomenie N, et al. Mitosis detection in breast cancer histological images an ICPR 2012 contest. J Pathol Inform 2013;4:8.
170. Spanhol FA, Oliveira LS, Petitjean C, Heutte L. A dataset for breast cancer histopathological image classification. IEEE Trans Biomed Eng 2016;63:1455–62.
171. Cruz-Roa A, Basavanhally A, Gonzalez F, et al. Automatic detection of invasive ductal carcinoma in whole slide images with Convolutional Neural Networks. In: Gurcan MN, Madabhushi A, editors. Medical Imaging 2014: Digital Pathology. Proceedings of SPIE. 90412014; 2014.
172. Araujo T, Aresta G, Castro E, et al. Classification of breast cancer histology images using Convolutional Neural Networks. PLoS ONE 2017;12:e0177544.
173. Dundar MM, Badve S, Bilgin G, et al. Computerized classification of intraductal breast lesions using histopathological images. IEEE Trans Biomed Eng 2011;58:1977–84.
174. Cruz-Roa A, Gilmore H, Basavanhally A, et al. Accurate and reproducible invasive breast cancer detection in whole-slide images: a Deep Learning approach for quantifying tumor extent. Sci Rep 2017;7:46450.
175. Litjens G, Sanchez CI, Timofeeva N, et al. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. Sci Rep 2016;6:26286.
176. Wang D, Khosla A, Gargeya R, Irshad H, Beck AH. Deep learning for identifying metastatic breast cancer. arXiv preprint arXiv:160605718. 2016.
177. Liu Y, Gadepalli K, Norouzi M, et al. Detecting cancer metastases on gigapixel pathology images. arXiv preprint arXiv:170302442. 2017.
178. Wollmann T, Rohr K. Automatic breast cancer grading in lymph nodes using a deep neural network. arXiv preprint arXiv: 170707565. 2017.
179. Vandenberghe ME, Scott ML, Scorer PW, Soderberg M, Balcerzak D, Barker C. Relevance of deep learning to facilitate the diagnosis of HER2 status in breast cancer. Sci Rep 2017;7: 45938.
180. Mungle T, Tewary S, Das DK, et al. MRF-ANN: a machine learning approach for automated ER scoring of breast cancer immunohistochemical images. J Microsc 2017;267:117–29.
181. Turkki R, Linder N, Kovanen PE, Pellinen T, Lundin J. Antibody-supervised deep learning for quantification of tumor-infiltrating immune cells in hematoxylin and eosin stained breast cancer samples. J Pathol Inform 2016;7:38.
182. Chen JM, Li Y, Xu J, et al. Computer-aided prognosis on breast cancer with hematoxylin and eosin histopathology images: a review. Tumour Biol 2017;39:1010428317694550.
183. Madabhushi A, Agner S, Basavanhally A, Doyle S, Lee G. Computer-aided prognosis: predicting patient and disease outcome via quantitative fusion of multi-scale, multi-modal data. Comput Med Imaging Graph 2011;35:506–14.
184. Beck AH, Sangoi AR, Leung S, et al. Systematic analysis of breast cancer morphology uncovers stromal features associated with survival. Sci Transl Med 2011;3:108ra13.
185. Szegedy C, Zaremba W, Sutskever I, et al. Intriguing properties of neural networks. arXiv preprint arXiv:13126199. 2013.
186. Xie P, Bilenko M, Finley T, Gilad-Bachrach R, Lauter K, Naehrig M. Crypto-nets: Neural networks over encrypted data. arXiv preprint arXiv:14126181. 2014.
187. Collins FS, Varmus H. A new initiative on precision medicine. N Engl J Med 2015;372:793–5.