

## Article

# Computer-Aided Assessment of Melanocytic Lesions by Means of a Mitosis Algorithm

Bart Sturm <sup>1,2</sup>, David Creytens <sup>3</sup>, Jan Smits <sup>2</sup>, Ariadne H. A. G. Ooms <sup>2</sup>, Erik Eijken <sup>4</sup>, Eline Kurpershoek <sup>2</sup>, Heidi V. N. Küsters-Vandeveld <sup>5</sup>, Carla Wauters <sup>5</sup>, Willeke A. M. Blokk <sup>6</sup> and Jeroen A. W. M. van der Laak <sup>1,7,\*</sup>

- <sup>1</sup> Department of Pathology, Radboud University Medical Center, 6500 HB Nijmegen, The Netherlands; b.sturm@pathan.nl
- <sup>2</sup> Pathan B.V., 3045 PM Rotterdam, The Netherlands; j.smits@pathan.nl (J.S.); a.ooms@pathan.nl (A.H.A.G.O.); e.kurpershoek@pathan.nl (E.K.)
- <sup>3</sup> Department of Pathology, Ghent University Hospital, 9000 Ghent, Belgium; david.creytens@uzgent.be
- <sup>4</sup> Laboratory for Pathology Oost Nederland (LabPON), 7550 AM Hengelo, The Netherlands; e.eijken@labpon.nl
- <sup>5</sup> Department of Pathology, Canisius Wilhelmina Hospital, 6500 GS Nijmegen, The Netherlands; h.kusters@cwz.nl (H.V.N.K.-V.); c.wauters@cwz.nl (C.W.)
- <sup>6</sup> Division Laboratories, Pharmacy and Biomedical Genetics, University Medical Center Utrecht, 3508 GA Utrecht, The Netherlands; W.A.M.Blokk@umcutrecht.nl
- <sup>7</sup> Center for Medical Image Science and Visualization, Linköping University, 581 83 Linköping, Sweden
- \* Correspondence: jeroen.vanderlaak@radboudumc.nl; Tel.: +31-638-814-869



**Citation:** Sturm, B.; Creytens, D.; Smits, J.; Ooms, A.H.A.G.; Eijken, E.; Kurpershoek, E.; Küsters-Vandeveld, H.V.N.; Wauters, C.; Blokk, W.A.M.; van der Laak, J.A.W.M. Computer-Aided Assessment of Melanocytic Lesions by Means of a Mitosis Algorithm. *Diagnostics* **2022**, *12*, 436. <https://doi.org/10.3390/diagnostics12020436>

Academic Editors: Chyi-Chia Richard Lee and Sung Chul Lim

Received: 26 November 2021

Accepted: 14 January 2022

Published: 8 February 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** An increasing number of pathology laboratories are now fully digitised, using whole slide imaging (WSI) for routine diagnostics. WSI paves the road to use artificial intelligence (AI) that will play an increasing role in computer-aided diagnosis (CAD). In melanocytic skin lesions, the presence of a dermal mitosis may be an important clue for an intermediate or a malignant lesion and may indicate worse prognosis. In this study a mitosis algorithm primarily developed for breast carcinoma is applied to melanocytic skin lesions. This study aimed to assess whether the algorithm could be used in diagnosing melanocytic lesions, and to study the added value in diagnosing melanocytic lesions in a practical setting. WSI's of a set of hematoxylin and eosin (H&E) stained slides of 99 melanocytic lesions (35 nevi, 4 intermediate melanocytic lesions, and 60 malignant melanomas, including 10 nevoid melanomas), for which a consensus diagnosis was reached by three academic pathologists, were subjected to a mitosis algorithm based on AI. Two academic and six general pathologists specialized in dermatopathology examined the WSI cases two times, first without mitosis annotations and after a washout period of at least 2 months with mitosis annotations based on the algorithm. The algorithm indicated true mitosis in lesional cells, i.e., melanocytes, and non-lesional cells, i.e., mainly keratinocytes and inflammatory cells. A high number of false positive mitosis was indicated as well, comprising melanin pigment, sebaceous glands nuclei, and spindle cell nuclei such as stromal cells and neuroid differentiated melanocytes. All but one pathologist reported more often a dermal mitosis with the mitosis algorithm, which on a regular basis, was incorrectly attributed to mitoses from mainly inflammatory cells. The overall concordance of the pathologists with the consensus diagnosis for all cases excluding nevoid melanoma ( $n = 89$ ) appeared to be comparable with and without the use of AI (89% vs. 90%). However, the concordance increased by using AI in nevoid melanoma cases ( $n = 10$ ) (75% vs. 68%). This study showed that in general cases, pathologists perform similarly with the aid of a mitosis algorithm developed primarily for breast cancer. In nevoid melanoma cases, pathologists perform better with the algorithm. From this study, it can be learned that pathologists need to be aware of potential pitfalls using CAD on H&E slides, e.g., misinterpreting dermal mitoses in non-melanotic cells.

**Keywords:** WSI; computer-aided diagnosis; mitosis algorithm; melanoma; nevoid melanoma

## 1. Introduction

Digital pathology is a dynamic, image-based environment that enables the acquisition, management, and interpretation of pathology information generated from a digitised glass slide, i.e., whole slide images (WSI), that can be assessed on a computer screen. Digital pathology offers all kinds of benefits, including digital archiving, consultation, and showcasing at tumour boards [1,2]. It is an innovation committed to the improvement of operational efficiency, including decreasing turn-around times with the reduction of laboratory expenses [2,3]. Since several WSI scanners are approved in Europe, given the European Conformity mark, in the United States of America by the Food and Drug Administration, and in Japan, by the Pharmaceuticals and Medical Devices Agency [4], enormous opportunities have arisen to analyse the sheer amount of slides with visual quantitative computer techniques, i.e., computational pathology (CP), based on machine learning (ML) [2]. CP may aid in assessing WSI for pathology diagnosis, so called computer-aided diagnosis (CAD). A range of different ML techniques are available of which in recent years, algorithms based on convolutional neural networks appear to dominate [5].

Malignant melanoma is the most lethal form of skin cancer and its prevalence varies among regions in the United States of America [6–9]. Research on CAD pathology has focused on different applications, including prostate cancer tumour detection and Gleason scoring [10–13] and breast cancer identification, grading (including assessing mitosis), hormone immunohistochemical status, and lymph node metastases [13–24]. Concerning CAD application in melanoma diagnostics, studies have been performed mainly on immunohistochemical (double) stains, e.g., phosphohistone H3, KI67, and/or MART1 [25,26]. Although diagnosing melanocytic lesion can be very challenging, CAD has not been studied extensively for melanoma pathology diagnostics on hematoxylin and eosin (H&E) stained slides. The presence and foremost the enumeration of subtle cytomorphologic and architectural features such as asymmetry of the lesion, cytological atypia, Pagetoid involvement of the epidermis, lack of maturation, presence of ulceration, and mitosis can make a major difference between a benign nevus, an intermediate lesion, or a malignant melanoma. Finding a mitosis in a melanotic cell, situated either epidermal or dermal, is of major importance as it implies the lesion might be intermediate or malignant [27]. It was shown that dermal mitoses indicate worse prognosis for survival and increased occurrence of sentinel node metastases [28–31].

In recent years, the development of computer-aided mitosis detection has increased significantly, which is partly due to publicly released training data sets for mitosis detection [32–35]. In a previous study, Tellez et al. trained a convolutional neural network (CNN) to detect individual mitotic figures in breast carcinoma WSI's with high accuracy [18]. The current study aimed to assess if a mitosis algorithm developed for breast cancer (1) can be used in the detection of mitosis in cutaneous melanocytic tumours, and (2) can improve the accuracy of the diagnosis of melanocytic lesions in a practical setting.

## 2. Materials and Methods

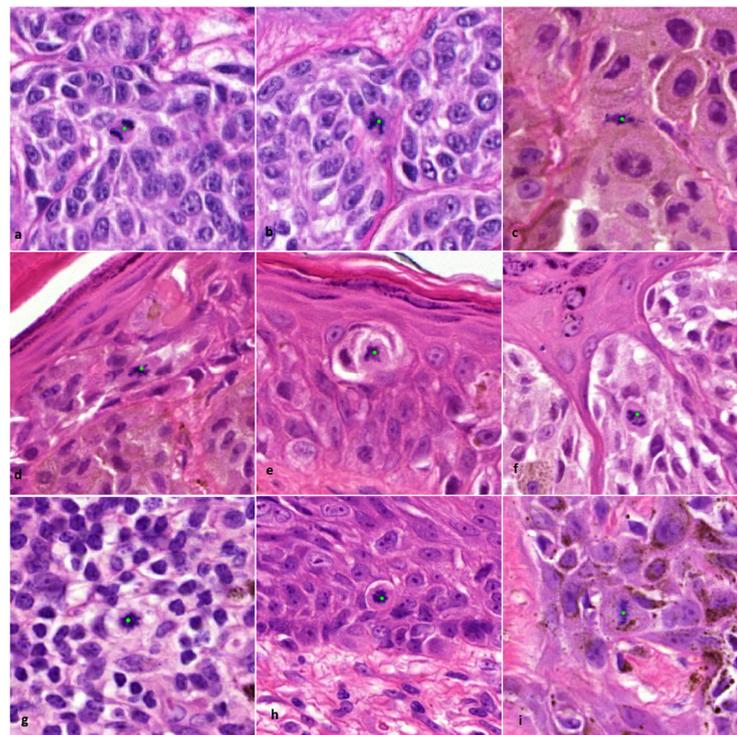
### 2.1. Case Selection and Study Design

In this study we used WSI of 102 H&E stained cases from a previous study, evaluating the potential added value of z-stack scanning in diagnosing melanocytic lesions [36]. The cases were obtained from the archive of the Pathology Department of the Radboud UMC in Nijmegen, The Netherlands, and concerned 35 benign nevi, 5 intermediate lesions (so-called melanocytomas or melanocytic tumours of unknown malignant potential; MELTUMP), and 62 malignant melanomas, including 10 nevoid melanomas. The set of WSI assessed by the study pathologists contained 99 cases for which consensus could be achieved by 3 academic pathologists based on the glass slides (35 benign melanocytic lesions, 4 intermediate cases, and 60 melanomas, including 10 nevoid melanomas) [36]. The cases, scanned with a Panoramic 250 flash II scanner (3D Histech, Budapest, Hungary), were re-randomized and submitted for evaluation to 8 pathologists that participated in the previous study (2 academic and 6 general pathologists). All WSI were assessed twice, first without and

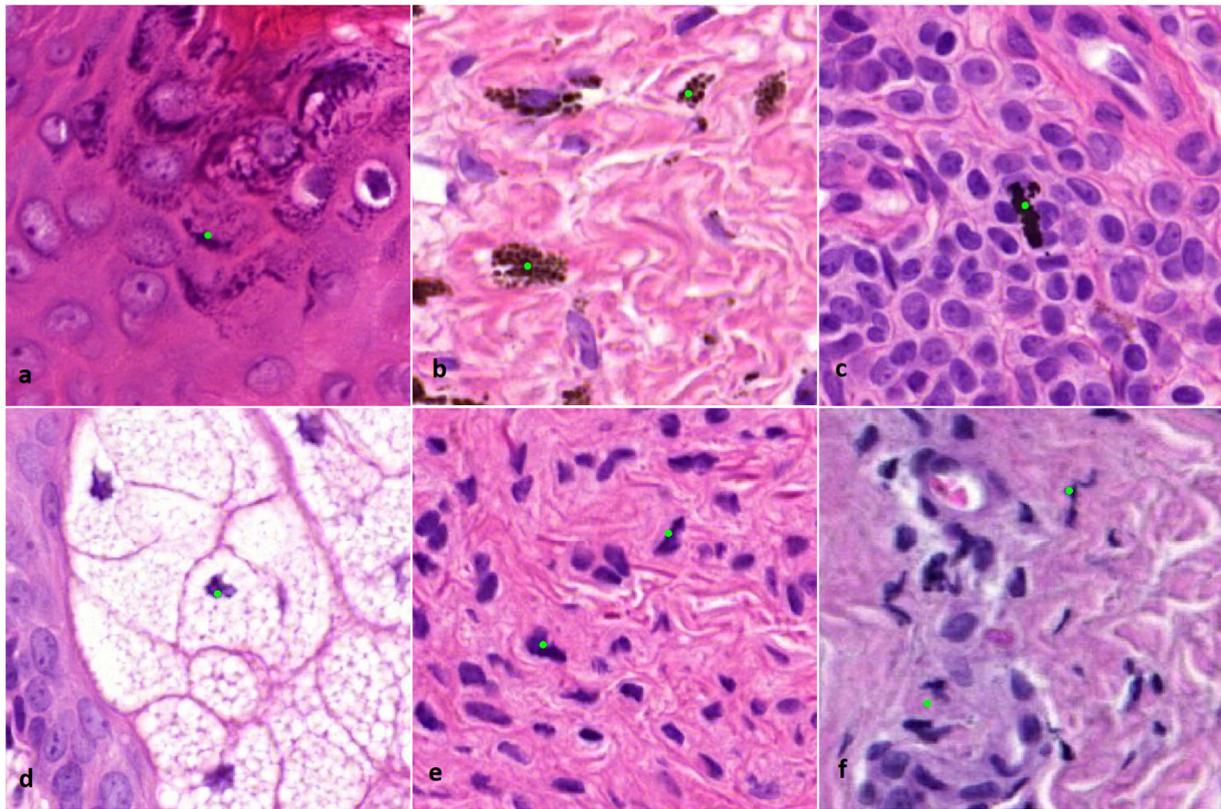
second with help of a mitosis detection algorithm, with a washout period of at least 2 months in between. The WSI cases were presented on a computer with a calibrated high resolution 4K LCD screen. Cases were offered with concise clinical information (age, gender, and location on the skin) and could be classified by the pathologists as either benign, malignant, or intermediate. For lesions classified as malignant or intermediate, the presence of dermal mitotic activity had to be reported. In addition, lesions classified as intermediate could be stratified into low risk and high risk. During the second assessment the pathologists were also asked to indicate which mitoses identified by the algorithm were helpful in making the diagnosis including cases that they classified as benign. More details on the study design are provided in our previous study [36].

## 2.2. Mitosis Algorithm

The mitosis algorithm used in this study was developed for automated detection of mitoses in breast carcinomas and is based on CNNs [18]. Therefore, in cutaneous melanocytic lesions false positive mitosis diagnoses were to be expected, because of different background stroma, colour, texture and the potential presence of melanin pigment. Before the start of the current study, a small pilot study was done to see if the algorithm was capable of detecting mitoses in 10 melanocytic cases. From these cases it was learned that mitoses could be identified correctly (Figure 1), although false positive mitoses were indicated by the algorithm as well (mostly consisting of melanin pigment, sebaceous gland nuclei, and spindle cell nuclei such as stromal cells and neuroid differentiated melanocytes) (Figure 2).



**Figure 1.** Examples of correctly indicated lesional and non-lesional mitoses by the algorithm ( $\times 800$  magnification). (a–c) Dermal mitosis in a melanocyte. (d–f) Epidermal mitosis in a melanocyte. (g) Dermal mitosis in an inflammatory cell. (h,i) Epidermal mitosis in a (pigmented) keratinocyte. Classification of indicated mitoses and type of cell origin was performed by an experienced pathologist (BS).



**Figure 2.** ( $\times 800$  magnification). Examples of incorrectly indicated mitoses by the algorithm. (a) Keratin granules. (b) Melanin pigment. (c) Formalin pigment. (d) Nucleus of a sebaceous cell. (e) Spindle cell nucleus of a melanocyte. (f) Squeezed nucleus of a lymphocyte.

In order to assess the algorithm's practical use in assisting pathologists to find mitoses in the current study, a selection of candidate mitoses was made (BS) before the annotated cases were offered to the participating pathologists. The selection excluded non-nucleated objects and nuclei of sebaceous glands specifically.

Application of the mitosis CNN was limited to a manually defined region of interest (ROI), using freely available ASAP software (version 1.8.1). The ROI was defined by an experienced pathologist (BS) and delineated the melanocytic lesion.

### 2.3. Statistical Analysis

For statistical analysis the four-tier scheme defined above was downsized to a three-tier system by combining high-risk and low-risk intermediate lesions, as discrimination of these lesions based on an H&E staining only without ancillary techniques is often not feasible [37]. Concordance of pathologists with the consensus diagnosis was expressed as the number and percentage of cases with identical diagnoses (in the three-tier system) for every subclass as well as overall. As an overall measure of concordance of the pathologists with the consensus diagnosis, Kappa statistics with 95% confidence intervals (CIs) were calculated. Data from our previous study was used for analysing the variation over time.

## 3. Results

In total, 2868 objects (range 1–676 per case) were detected by the algorithm in 76 cases that classified for a (candidate) mitosis. After manual selection, 825 (candidate) mitoses were retained that were annotated in 61 cases comprising mainly epidermal and dermal mitoses, i.e., lesional but also non-lesional mitoses in keratinocytes and inflammatory cells. Furthermore, spindle cell nuclei, apoptotic cells and contused nuclei of e.g., lymphocytes, were annotated as well.

The overall concordance of the pathologists with the consensus diagnosis for all cases excluding nevoid melanoma ( $n = 89$ ) appears to be comparable with and without the use of AI (89% vs. 90%), shown in Table 1. Agreement according to Cohen’s Kappa is at least substantial with and without the mitosis algorithm, shown in Table 2, except for pathologist PATH4 reaching at least moderate agreement with the consensus diagnosis.

**Table 1.** Concordance as a percentage (%) of cases identical with the consensus diagnosis based on glass slides for all cases excluding nevoid melanoma ( $n = 89$ ) and number of cases with reported dermal mitosis (#DM) concerning intermediate lesions and melanoma (excl. nevoid melanoma) ( $n = 54$ ). <sup>a</sup> Glass; <sup>b</sup> WSI.

Pathologist	z-Stack Study [36] Glass <sup>a</sup> /WSI <sup>b</sup>		1st Round WSI		2nd Round WSI Algorithm	
	%	#DM	%	#DM	%	#DM
EXP1	97 <sup>a</sup>	30 <sup>a</sup>	94	27	91	27
EXP2	93 <sup>a</sup>	22 <sup>a</sup>	91	26	89	27
PATH1	89 <sup>b</sup>	19 <sup>b</sup>	91	25	91	44
PATH2	89 <sup>b</sup>	22 <sup>b</sup>	96	25	94	35
PATH3	81 <sup>b</sup>	20 <sup>b</sup>	87	26	92	27
PATH4	75 <sup>b</sup>	17 <sup>b</sup>	84	27	76	40
PATH5	84 <sup>b</sup>	21 <sup>b</sup>	96	25	94	29
PATH6	90 <sup>b</sup>	18 <sup>b</sup>	84	21	84	21
Average	87 <sup>a,b</sup>	21 <sup>a,b</sup>	90	25	89	31

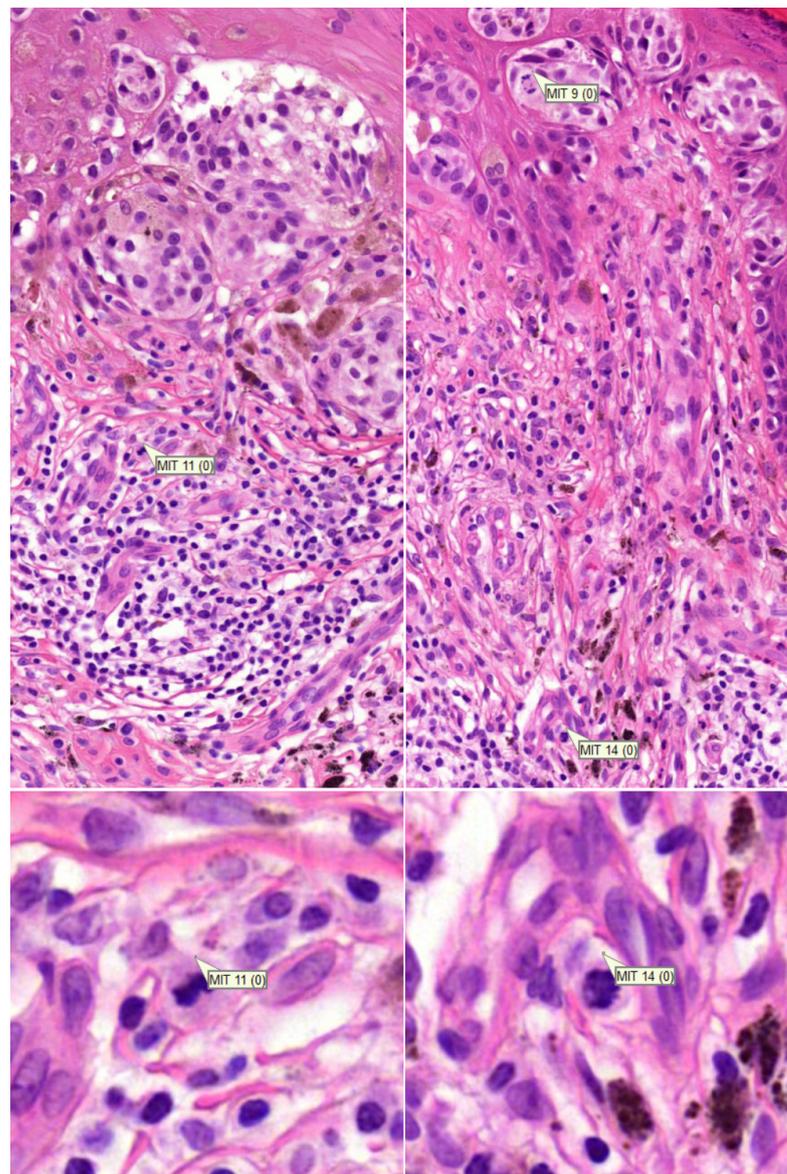
According to the consensus diagnosis, dermal mitoses are present in 28 cases. The table is intended to give an overview of the concordance and number of reported dermal mitoses over time and doesn’t show if the reported dermal mitoses are correct with respect to the consensus diagnosis.

**Table 2.** Kappa values (95% confidence interval) for all cases excluding nevoid melanoma ( $n = 89$ ). <sup>a</sup> Glass; <sup>b</sup> WSI.

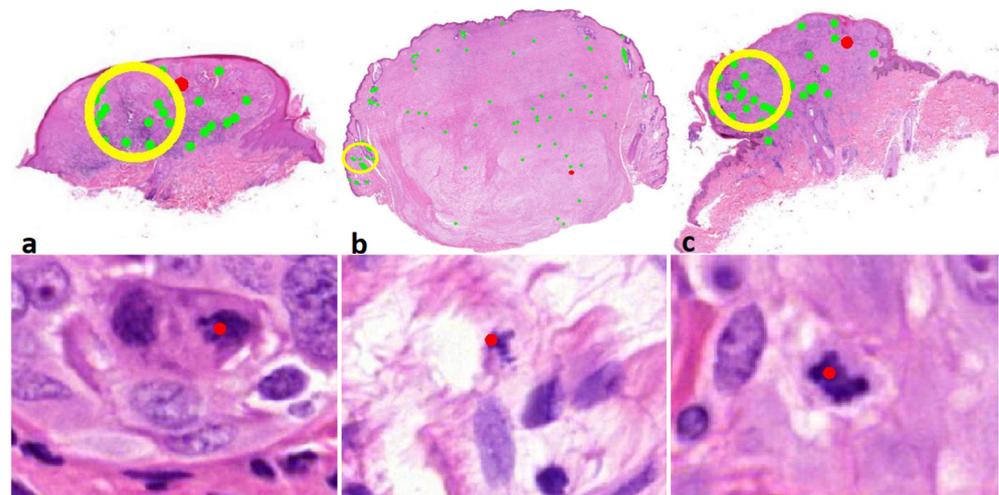
Pathologist	z-Stack Study Glass <sup>a</sup> /WSI <sup>b</sup>	1st Round WSI	2nd Round WSI Algorithm
EXP1	0.94 <sup>a</sup>	0.89 (0.80–0.98)	0.83 (0.73–0.94)
EXP2	0.88 <sup>a</sup>	0.83 (0.72–0.94)	0.79 (0.67–0.91)
PATH1	0.78 (0.66–0.90) <sup>b</sup>	0.83 (0.72–0.94)	0.83 (0.72–0.94)
PATH2	0.79 (0.67–0.91) <sup>b</sup>	0.92 (0.84–1.00)	0.89 (0.80–0.98)
PATH3	0.66 (0.51–0.79) <sup>b</sup>	0.76 (0.64–0.88)	0.85 (0.75–0.95)
PATH4	0.55 (0.39–0.70) <sup>b</sup>	0.70 (0.56–0.84)	0.55 (0.41–0.69)
PATH5	0.72 (0.58–0.83) <sup>b</sup>	0.92 (0.84–1.00)	0.90 (0.81–0.98)
PATH6	0.81 (0.69–0.91) <sup>b</sup>	0.73 (0.61–0.85)	0.73 (0.61–0.85)

In Table 1 the number of cases with reported dermal mitoses are presented as well. According to the consensus diagnosis, in 28 cases out of a total of 54 malignant (excl. nevoid melanoma) plus intermediate cases, dermal mitoses were present. All pathologists excluding one academic pathologist reported more dermal mitoses with the mitosis algorithm. Three pathologists reported a substantially higher number of cases with dermal mitoses (PATH1, PATH2, and PATH4). After reviewing these cases (BS), it appeared that on a regular basis, mitoses in infiltrates nearby the tumour front were interpreted as falsely being mitoses from melanocytes (Figure 3). However, in three cases (cases 56, 73, and 100), at least one pathologist reported a dermal mitosis by means of the algorithm that was reconfirmed by the investigator (BS), and was discordant with the consensus diagnosis

based on glass slides (Figure 4), although the consensus concerning the presence of a dermal mitosis was not unanimous at the time. Furthermore, pathologist PATH4 had a significant lower concordance rate with the mitosis algorithm, where the algorithm did aid in correctly changing the diagnosis from benign to malignant in six cases versus incorrectly changing the diagnosis four times. The lower concordance was mainly due to incorrectly changing the diagnosis from benign to intermediate in eight cases in which two cases of dermal mitosis was reported incorrectly with the mitosis algorithm.



**Figure 3.** (row above  $\times 200$ , row below  $\times 800$  magnification). Examples of incorrectly attributed mitoses from inflammatory cells located outside of the tumour front.

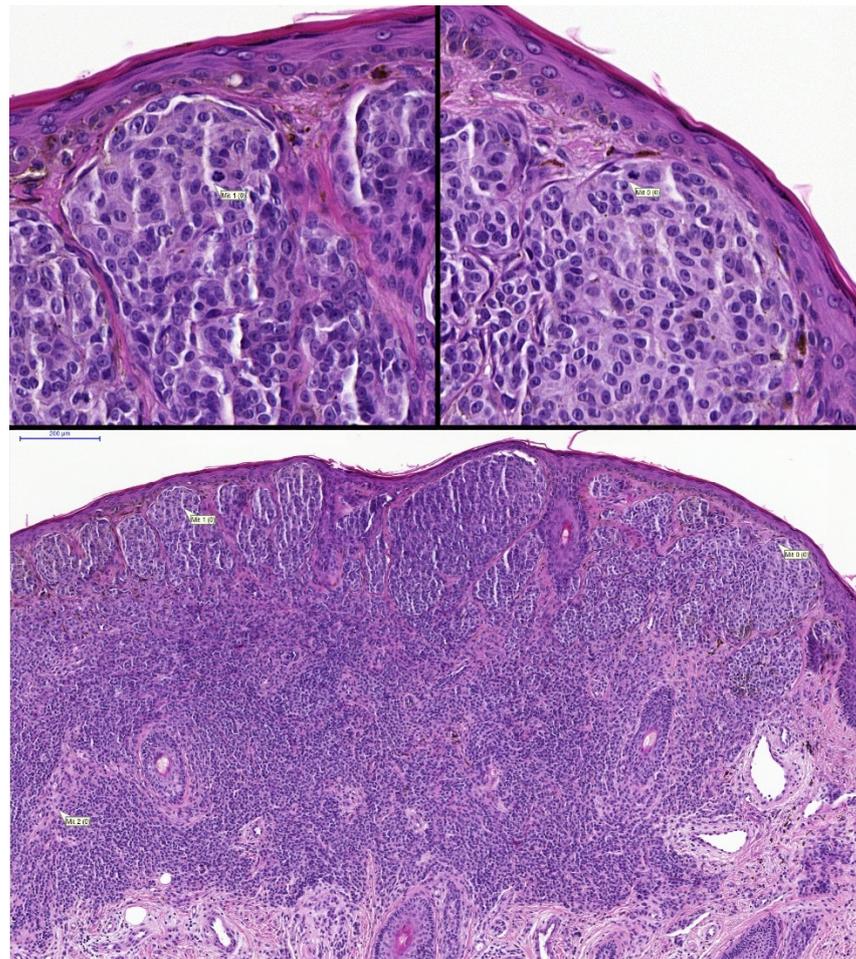


**Figure 4.** (row above  $\times 6$ , row below  $\times 800$  magnification). Correct dermal mitosis discordant with the consensus. Above overview with red indicates mitosis, below a detailed view of the indicated mitosis. (a) Case 56; melanocytoma/intermediate lesion, high risk. (b) Case 73; desmoplastic melanoma. (c) Case 100; melanocytoma/intermediate lesion, low risk.

In general, the pathologists appeared to have an advantage with the mitosis algorithm in the nevoid melanoma cases ( $n = 10$ ), as shown in Table 3 and example given in Figure 5, although the number of cases with reported dermal mitosis on average kept constant with and without the mitosis algorithm, i.e., five to six. If melanoma and intermediate diagnosis were grouped together, all but one pathologist performed better with the algorithm, recognizing the nevoid melanoma cases as being at least potentially malignant.

**Table 3.** Concordance of nevoid melanoma ( $n = 10$ ) as a percentage (%) of cases identical with the consensus diagnosis based on glass slides and number of cases with reported dermal mitosis (#DM).

Pathologist	z-Stack Study WSI		1st Round WSI		2nd Round WSI Algorithm	
	%	#DM	%	#DM	%	#DM
EXP1	-	-	70	7	70	5
EXP2	-	-	80	4	90	4
PATH1	50	4	70	6	80	8
PATH2	70	6	90	9	90	8
PATH3	20	0	40	3	70	3
PATH4	10	1	50	7	90	7
PATH5	80	4	70	5	70	6
PATH6	80	3	70	4	40	3
Average	52	3	68	5,6	75	5,5



**Figure 5.** (row above  $\times 400$ , row below  $\times 100$  magnification). Example of nevoid melanoma with two annotated correct lesional dermal mitoses. Case 71; two pathologists changed the diagnosis from benign to malignant with the aid of the algorithm.

#### 4. Discussion

The application of AI in routine pathology diagnostics is on the rise. This study is the first, to the best of our knowledge, to actually use a CNN-based mitosis algorithm to aid pathologists in assessing melanocytic lesions in a routine diagnostic setting. Former studies have shown that a deep learning algorithm has the potential to improve diagnostic workflow in diagnosis i.e., nodular basal cell carcinoma, seborrheic keratosis, dermal nevus, and melanoma [38,39]. Andres et al. presented a proof-of-principle of a computer-aided staging support system for malignant melanoma [40]. Studies have shown that immunohistochemical markers i.e., PHH3, Ki-67, P16, HM45, and PRAME can aid pathologists in the diagnosis of melanoma and may reduce observer variation [26,41–45]. However, for pathologists it is more convenient to make all analyses, including CAD, on H&E WSI, mostly because it is less time-consuming and expected to be less expensive than immunohistochemistry in the coming years. Mitosis detection on H&E slides has been investigated thoroughly for breast cancer histopathology, where mitotic density was shown to be prognostic and is therefore part of the grading system and is of importance for patient management [19]. For melanoma diagnosis, the performances of mitosis algorithms have been tested in skin tissue with promising results [40,46]. Studies investigating tumour grading have reported that it is difficult to establish a ground truth for mitotic cells. Criteria for defining a mitosis can be given but in practice these criteria are difficult to follow in a strict manner by pathologists because of doubtful instances where subjective interpretations must be made. Misinterpretation of mitoses can occur due to similarity to, for example, apoptotic cells. As a result,

a relatively large inter- and intra-observer variation in recognizing mitoses is a fact [19]. This is reflected in our previous study on the effect of z-stack scanning, where pathologists reported dermal mitoses in a range from 17 to 30 cases ( $n = 54$ ) [36]. Nevertheless, in the current study, even without the algorithm, the range narrowed to 21–27. This may be a result of more experience with assessing WSI's by the pathologists, since the execution of the former study was during the years 2016 and 2017.

Tabata et al. found that traditional light microscopy was measurably more accurate in detecting mitoses than WSI [4]. This may be explained because of the omission of focussing in different z-planes in WSI. Another explanation is that WSI's are scanned on 20× objective while in microscopy assessment mitoses are mostly found on 40× objective. Although it is reported that in WSI less mitoses are detected compared to light microscopy, in our study a larger amount of dermal mitoses was reported on WSI with the aid of the mitosis algorithm. On a regular basis, these mitoses were attributed falsely by some pathologists to mitoses of inflammatory cells and, to a lesser extent, to the difficulty of interpreting mitoses as, for example in squeezed nuclei of lymphocytes, bizarre nuclei of melanomas or apoptotic cells. However, in three cases at least one pathologist reported a dermal mitosis by means of the algorithm that was discordant with the ground truth based on glass slides underpinning the potential advantage of the algorithm.

Concerning the nevoid melanoma cases, a mitosis was found sporadically with the aid of the algorithm, changing the diagnosis from benign to either intermediate or melanoma. Remarkable is the fact that only two of the eight pathologists reported more dermal mitosis in the nevoid melanoma cases with the mitosis algorithm, although it did not aid significantly in recognizing nevoid melanoma. In case 16 and 71, respectively, three and two pathologists changed their diagnosis from benign to malignant. In these two cases pathologists reported, respectively, one and four dermal mitoses to be of aid for the diagnosis. After reviewing the reported mitoses these were dermal mitoses in melanocytes (BS). On the other hand, in some nevoid melanoma cases, despite the awareness of the presence of dermal mitoses by the pathologist, these melanocytic lesions were still falsely interpreted by several pathologists as benign, reflecting the complex interpretation of this class of malignant lesions.

Finally, in this study, a high number of false positive mitotic objects were indicated by the algorithm, e.g., spindle cell nuclei of stromal and melanocytic cells and squeezed nuclei of mainly inflammatory cells, which appeared time-consuming for the pathologists to assess. These false positive objects were an expected finding as the CNN was not trained for skin tissue and melanocytic tumours, mainly due to a shortage of resources to tune the algorithm for this purpose. Although this is a limitation of our study design, we did overcome this by a manual selection of (candidate) mitoses. Nevertheless, a CNN can be optimized in finding mitoses while simultaneously discriminating similar objects that are not of interest to the pathologist. A perfect mitosis algorithm will indicate all mitoses, not limited to the cell of interest, in the study of the melanocyte. This points out another limitation of the study, that the mitosis algorithm doesn't discriminate between a mitosis from a melanocyte, keratinocyte, lymphocyte, or other cell. Immunohistochemical double stains (PHH3, Ki-67 and/or MART1) can overcome this uncertainty, although it is costly and time-consuming. In daily practice, the cell of origin is mainly classified on the basis of the location of the mitosis, i.e., in a lesional or non-lesional area. In order to effectively aid pathologists in identifying mitosis in melanocytic lesions, an algorithm should preferentially discriminate mitosis in a melanocytic cell and omit mitoses mainly from keratinocytes and inflammatory cells, as well as difficult-to-interpret objects such as, for example, squeezed nuclei. Therefore, optimization in the differentiation of mitosis (like) objects may be one of the next objectives in CNN mitosis detection development, e.g., to discriminate areas of interest in lesional epidermis, lesional dermis, and non-lesional stroma.

## 5. Conclusions

Diagnosing melanocytic lesions is challenging and may have major implications on patient management and wellbeing. This study shows that a mitosis algorithm that was primarily developed for breast cancer can be applied to melanocytic skin lesions, although it is not applicable in a practical setting due to a high number of false positive-indicated mitoses. After a correction procedure for false positive (candidate) mitoses, it appeared that in general cases, pathologists perform similarly with the aid of a mitosis algorithm in WSI. However, pathologists perform better with the algorithm in nevoid melanoma cases, which are notoriously difficult to recognise. From this study it can be learned that pathologists need to be aware of potential pitfalls using computer-aided diagnosis on H&E slides, such as misinterpreting dermal mitosis from non-melanotic cells, i.e., mainly inflammatory cells.

**Author Contributions:** Conceptualization, B.S. and J.A.W.M.v.d.L.; methodology, B.S. and J.A.W.M.v.d.L.; software, B.S. and J.A.W.M.v.d.L.; validation, B.S. and J.A.W.M.v.d.L.; formal analysis, B.S. and J.A.W.M.v.d.L.; investigation, B.S., J.A.W.M.v.d.L., D.C., J.S., A.H.A.G.O., E.E., E.K., H.V.N.K.-V., C.W. and W.A.M.B.; resources, B.S. and J.A.W.M.v.d.L.; data curation, B.S. and J.A.W.M.v.d.L.; writing—original draft preparation, B.S.; writing—review & editing: J.A.W.M.v.d.L., W.A.M.B., A.H.A.G.O., H.V.N.K.-V. and C.W.; visualization, B.S. and J.A.W.M.v.d.L.; supervision, J.A.W.M.v.d.L. and W.A.M.B.; project administration, B.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This project received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 945358. This Joint Undertaking receives support from the European Union’s Horizon 2020 research and innovation program and EFPIA.

**Acknowledgments:** I would like to thank Caner Mercan and Maschenka Balkenhol for running the mitosis algorithm analysis on the whole slide images and collecting the data. The Knut and Alice Wallenberg foundation is acknowledged for their generous support.

**Conflicts of Interest:** Jeroen van der Laak was a member of the advisory boards of Philips, the Netherlands and ContextVision, Sweden, and received research funding from Philips, the Netherlands, ContextVision, Sweden, and Sectra, Sweden in the last five years. He is chief scientific officer (CSO) and shareholder of Aiosyn BV, the Netherlands. The other authors declare no conflict of interest.

## References

- Farahani, N.; Parwani, A.V.; Pantanowitz, L. Whole slide imaging in pathology: Advantages, limitations, and emerging perspectives. *Pathol. Lab. Med. Int.* **2015**, *7*, 23–33. [[CrossRef](#)]
- Tomaszewski, J. Overview of the role of artificial intelligence in pathology: The computer as a pathology digital assistant. In *Artificial Intelligence and Deep Learning in Pathology*, 1st ed.; Cohen, S., Ed.; Elsevier: Amsterdam, The Netherlands, 2021.
- Baidoshvili, A.; Bucur, A.; van Leeuwen, J.; van der Laak, J.; Kluin, P.; van Diest, P.J. Evaluating the benefits of digital pathology implementation: Time savings in laboratory logistics. *Histopathology* **2018**, *73*, 784–794. [[CrossRef](#)] [[PubMed](#)]
- Tabata, K.; Uraoka, N.; Benhamida, J.; Hanna, M.G.; Sirintrapun, S.J.; Gallas, B.D.; Gong, Q.; Aly, R.G.; Emoto, K.; Matsuda, K.M.; et al. Validation of mitotic cell quantification via microscopy and multiple whole-slide scanners. *Diagn. Pathol.* **2019**, *14*, 65. [[CrossRef](#)] [[PubMed](#)]
- Cohen, S. The evolution of machine learning: Past, present, and future. In *Artificial Intelligence and Deep Learning in Pathology*, 1st ed.; Cohen, S., Ed.; Elsevier: Amsterdam, The Netherlands, 2021.
- Linos, E.; Swetter, S.M.; Cockburn, M.G.; Colditz, G.A.; Clarke, C.A. Increasing burden of melanoma in the United States. *J. Investig. Dermatol.* **2009**, *129*, 1666–1674. [[CrossRef](#)] [[PubMed](#)]
- Howlander, N.; Noone, A.M.; Krapcho, M.; Miller, D.; Brest, A.; Yu, M.; Ruhl, J.; Tatalovich, Z.; Mariotto, A.; Lewis, D.R.; et al. (Eds.) *SEER Cancer Statistics Review, 1975–2018*; National Cancer Institute: Bethesda, MD, USA, 2021.
- Erdei, E.; Torres, S.M. A new understanding in the epidemiology of melanoma. *Expert Rev. Anticancer* **2010**, *10*, 1811–1823. [[CrossRef](#)]
- Laughter, M.R.; Maymone, M.B.C.; Karimkhani, C.; Rundle, C.; Hu, S.; Wolfe, S.; Abuabara, K.; Hollingsworth, P.; Weintraub, G.S.; Dunnick, C.A.; et al. The Burden of Skin and Subcutaneous Diseases in the United States from 1990 to 2017. *JAMA Dermatol.* **2020**, *156*, 874–881. [[CrossRef](#)] [[PubMed](#)]
- Bulten, W.; Pinckaers, H.; van Boven, H.; Vink, R.; de Bel, T.; van Ginneken, B.; van der Laak, J.; Hulsbergen-van de Kaa, C.; Litjens, G. Automated deep-learning system for Gleason grading of prostate cancer using biopsies: A diagnostic study. *Lancet Oncol.* **2020**, *21*, 233–241. [[CrossRef](#)]
- Tabesh, A.; Teverovskiy, M.; Pang, H.Y.; Kumar, V.P.; Verbel, D.; Kotsianti, A.; Saidi, O. Multifeature prostate cancer diagnosis and Gleason grading of histological images. *IEEE Trans. Med. Imaging* **2007**, *26*, 1366–1378. [[CrossRef](#)]

12. Han, W.; Johnson, C.; Gaed, M.; Gomez, J.A.; Moussa, M.; Chin, J.L.; Pautler, S.; Bauman, G.S.; Ward, A.D. Histologic tissue components provide major cues for machine learning-based prostate cancer detection and grading on prostatectomy specimens. *Sci. Rep.* **2020**, *10*, 9911. [[CrossRef](#)]
13. Litjens, G.; Sanchez, C.I.; Timofeeva, N.; Hermsen, M.; Nagtegaal, I.; Kovacs, I.; Hulsbergen-van de Kaa, C.; Bult, P.; van Ginneken, B.; van der Laak, J. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Sci. Rep.* **2016**, *6*, 26286. [[CrossRef](#)]
14. Keay, T.; Conway, C.M.; O'Flaherty, N.; Hewitt, S.M.; Shea, K.; Gavrielides, M.A. Reproducibility in the automated quantitative assessment of HER2/neu for breast cancer. *J. Pathol. Inf.* **2013**, *4*, 19. [[CrossRef](#)]
15. Gavrielides, M.A.; Gallas, B.D.; Lenz, P.; Badano, A.; Hewitt, S.M. Observer variability in the interpretation of HER2/neu immunohistochemical expression with unaided and computer-aided digital microscopy. *Arch. Pathol. Lab. Med.* **2011**, *135*, 233–242. [[CrossRef](#)] [[PubMed](#)]
16. Ehteshami Bejnordi, B.; Veta, M.; Johannes van Diest, P.; van Ginneken, B.; Karssemeijer, N.; Litjens, G.; van der Laak, J.; Hermsen, M.; Manson, Q.F.; Balkenhol, M.; et al. Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women with Breast Cancer. *JAMA* **2017**, *318*, 2199–2210. [[CrossRef](#)] [[PubMed](#)]
17. Steiner, D.F.; MacDonald, R.; Liu, Y.; Truszkowski, P.; Hipp, J.D.; Gammage, C.; Thng, F.; Peng, L.; Stumpe, M.C. Impact of Deep Learning Assistance on the Histopathologic Review of Lymph Nodes for Metastatic Breast Cancer. *Am. J. Surg. Pathol.* **2018**, *42*, 1636–1646. [[CrossRef](#)]
18. Tellez, D.; Balkenhol, M.; Otte-Holler, I.; van de Loo, R.; Vogels, R.; Bult, P.; Wauters, C.; Vreuls, W.; Mol, S.; Karssemeijer, N.; et al. Whole-Slide Mitosis Detection in H&E Breast Histology Using PHH3 as a Reference to Train Distilled Stain-Invariant Convolutional Networks. *IEEE Trans. Med. Imaging* **2018**, *37*, 2126–2136. [[CrossRef](#)]
19. Veta, M.; van Diest, P.J.; Jiwa, M.; Al-Janabi, S.; Pluim, J.P. Mitosis Counting in Breast Cancer: Object-Level Interobserver Agreement and Comparison to an Automatic Method. *PLoS ONE* **2016**, *11*, e0161286. [[CrossRef](#)]
20. Balkenhol, M.C.A.; Tellez, D.; Vreuls, W.; Clahsen, P.C.; Pinckaers, H.; Ciompi, F.; Bult, P.; van der Laak, J. Deep learning assisted mitotic counting for breast cancer. *Lab. Investig.* **2019**, *99*, 1596–1606. [[CrossRef](#)]
21. Saha, M.; Chakraborty, C.; Racoceanu, D. Efficient deep learning model for mitosis detection using breast histopathology images. *Comput. Med. Imaging Graph.* **2018**, *64*, 29–40. [[CrossRef](#)]
22. Bejnordi, B.E.; Zuidhof, G.; Balkenhol, M.; Hermsen, M.; Bult, P.; van Ginneken, B.; Karssemeijer, N.; Litjens, G.; van der Laak, J. Context-aware stacked convolutional neural networks for classification of breast carcinomas in whole-slide histopathology images. *J. Med. Imaging* **2017**, *4*, 044504. [[CrossRef](#)]
23. Cruz-Roa, A.; Gilmore, H.; Basavanahally, A.; Feldman, M.; Ganesan, S.; Shih, N.N.C.; Tomaszewski, J.; Gonzalez, F.A.; Madabhushi, A. Accurate and reproducible invasive breast cancer detection in whole-slide images: A Deep Learning approach for quantifying tumor extent. *Sci. Rep.* **2017**, *7*, 46450. [[CrossRef](#)] [[PubMed](#)]
24. Araujo, T.; Aresta, G.; Castro, E.; Rouco, J.; Aguiar, P.; Eloy, C.; Polonia, A.; Campilho, A. Classification of breast cancer histology images using Convolutional Neural Networks. *PLoS ONE* **2017**, *12*, e0177544. [[CrossRef](#)] [[PubMed](#)]
25. Nielsen, P.S.; Riber-Hansen, R.; Jensen, T.O.; 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*, 404–413. [[CrossRef](#)] [[PubMed](#)]
26. Wandler, A.; Spaun, E.; Steiniche, T.; Nielsen, P.S. Automated quantification of Ki67/MART1 stains may prevent false-negative melanoma diagnoses. *J. Cutan. Pathol.* **2016**, *43*, 956–962. [[CrossRef](#)]
27. Filosa, A.; Filosa, G. Melanoma Diagnosis: The Importance of Histopathological Report. *Dermatopathology* **2018**, *5*, 41–43. [[CrossRef](#)]
28. Francken, A.B.; Shaw, H.M.; Thompson, J.F.; Soong, S.J.; Accortt, N.A.; Azzola, M.F.; Scolyer, R.A.; Milton, G.W.; McCarthy, W.H.; Colman, M.H.; et al. The prognostic importance of tumor mitotic rate confirmed in 1317 patients with primary cutaneous melanoma and long follow-up. *Ann. Surg. Oncol.* **2004**, *11*, 426–433. [[CrossRef](#)]
29. Barnhill, R.L.; Katzen, J.; Spatz, A.; Fine, J.; Berwick, M. The importance of mitotic rate as a prognostic factor for localized cutaneous melanoma. *J. Cutan. Pathol.* **2005**, *32*, 268–273. [[CrossRef](#)] [[PubMed](#)]
30. Nagore, E.; Oliver, V.; Botella-Estrada, R.; Moreno-Picot, S.; Insa, A.; Fortea, J.M. Prognostic factors in localized invasive cutaneous melanoma: High value of mitotic rate, vascular invasion and microscopic satellitosis. *Melanoma Res.* **2005**, *15*, 169–177. [[CrossRef](#)]
31. Conic, R.R.Z.; Ko, J.; Damiani, G.; Funchain, P.; Knackstedt, T.; Vij, A.; Vidimos, A.; Gastman, B.R. Predictors of sentinel lymph node positivity in thin melanoma using the National Cancer Database. *J. Am. Acad. Dermatol.* **2019**, *80*, 441–447. [[CrossRef](#)]
32. Roux, L.; Racoceanu, D.; Lomenie, N.; Kulikova, M.; Irshad, H.; Klossa, J.; Capron, F.; Genestie, C.; Le Naour, G.; Gurcan, M.N. Mitosis detection in breast cancer histological images an ICPR 2012 contest. *J. Pathol. Inf.* **2013**, *4*, 8. [[CrossRef](#)] [[PubMed](#)]
33. Veta, M.; van Diest, P.J.; Willems, S.M.; Wang, H.; Madabhushi, A.; Cruz-Roa, A.; Gonzalez, F.; Larsen, A.B.; Vestergaard, J.S.; Dahl, A.B.; et al. Assessment of algorithms for mitosis detection in breast cancer histopathology images. *Med. Image Anal.* **2015**, *20*, 237–248. [[CrossRef](#)]
34. Roux, L. MITOS-ATYPIA-14. Available online: <https://mitos-atypia-14.grand-challenge.org> (accessed on 12 May 2021).
35. Veta, M.; Heng, Y.J.; Stathonikos, N.; Bejnordi, B.E.; Beca, F.; Wollmann, T.; Rohr, K.; Shah, M.A.; Wang, D.; Rousson, M.; et al. Predicting breast tumor proliferation from whole-slide images: The TUPAC16 challenge. *Med. Image Anal.* **2019**, *54*, 111–121. [[CrossRef](#)] [[PubMed](#)]

36. Sturm, B.; Creyten, D.; Cook, M.G.; Smits, J.; van Dijk, M.; Eijken, E.; Kurpershoek, E.; Kusters-Vandeveld, H.V.N.; Ooms, A.; Wauters, C.; et al. Validation of Whole-slide Digitally Imaged Melanocytic Lesions: Does Z-Stack Scanning Improve Diagnostic Accuracy? *J. Pathol. Inf.* **2019**, *10*, 6. [[CrossRef](#)]
37. De la Fouchardiere, A.; Blokk, W.; van Kempen, L.C.; Luzar, B.; Piperno-Neumann, S.; Puig, S.; Alos, L.; Calonje, E.; Massi, D.; Group, E.S.P.D.W.; et al. ESP, EORTC, and EURACAN Expert Opinion: Practical recommendations for the pathological diagnosis and clinical management of intermediate melanocytic tumors and rare related melanoma variants. *Virchows Arch.* **2021**, *479*, 3–11. [[CrossRef](#)]
38. Olsen, T.G.; Jackson, B.H.; Feeser, T.A.; Kent, M.N.; Moad, J.C.; Krishnamurthy, S.; Lunsford, D.D.; Soans, R.E. Diagnostic Performance of Deep Learning Algorithms Applied to Three Common Diagnoses in Dermatopathology. *J. Pathol. Inf.* **2018**, *9*, 32. [[CrossRef](#)]
39. Hekler, A.; Utikal, J.S.; Enk, A.H.; Berking, C.; Klode, J.; Schadendorf, D.; Jansen, P.; Franklin, C.; Holland-Letz, T.; Krahl, D.; et al. Pathologist-level classification of histopathological melanoma images with deep neural networks. *Eur. J. Cancer* **2019**, *115*, 79–83. [[CrossRef](#)]
40. Andres, C.; Andres-Belloni, B.; Hein, R.; Biedermann, T.; Schape, A.; Brieu, N.; Schonmeyer, R.; Yigitsoy, M.; Ring, J.; Schmidt, G.; et al. iDermatoPath—A novel software tool for mitosis detection in H&E-stained tissue sections of malignant melanoma. *J. Eur. Acad. Dermatol. Venereol.* **2017**, *31*, 1137–1147. [[CrossRef](#)]
41. Nasr, M.R.; El-Zammar, O. Comparison of pHH3, Ki-67, and survivin immunoreactivity in benign and malignant melanocytic lesions. *Am. J. Dermatol.* **2008**, *30*, 117–122. [[CrossRef](#)]
42. Uguen, A.; Talagas, M.; Costa, S.; Duigou, S.; Bouvier, S.; De Braekeleer, M.; Marcorelles, P. A p16-Ki-67-HMB45 immunohistochemistry scoring system as an ancillary diagnostic tool in the diagnosis of melanoma. *Diagn. Pathol.* **2015**, *10*, 195. [[CrossRef](#)]
43. Lezcano, C.; Jungbluth, A.A.; Nehal, K.S.; Hollmann, T.J.; Busam, K.J. PRAME Expression in Melanocytic Tumors. *Am. J. Surg. Pathol.* **2018**, *42*, 1456–1465. [[CrossRef](#)] [[PubMed](#)]
44. Gassenmaier, M.; Hahn, M.; Metzler, G.; Bauer, J.; Yazdi, A.S.; Keim, U.; Garbe, C.; Wagner, N.B.; Forchhammer, S. Diffuse PRAME Expression Is Highly Specific for Thin Melanomas in the Distinction from Severely Dysplastic Nevi but Does Not Distinguish Metastasizing from Non-Metastasizing Thin Melanomas. *Cancers* **2021**, *13*, 3864. [[CrossRef](#)]
45. Lohman, M.E.; Steen, A.J.; Grekin, R.C.; North, J.P. The utility of PRAME staining in identifying malignant transformation of melanocytic nevi. *J. Cutan. Pathol.* **2021**, *48*, 856–862. [[CrossRef](#)]
46. Nofallah, S.; Mehta, S.; Mercan, E.; Knezevich, S.; May, C.J.; Weaver, D.; Witten, D.; Elmore, J.G.; Shapiro, L. Machine learning techniques for mitoses classification. *Comput. Med. Imaging Graph.* **2021**, *87*, 101832. [[CrossRef](#)]