


RESEARCH ARTICLE

Open Access



# Leveraging artificial intelligence to predict ERG gene fusion status in prostate cancer

Vipulkumar Dadhania<sup>1</sup>, Daniel Gonzalez<sup>2</sup>, Mustafa Yousif<sup>1,3</sup>, Jerome Cheng<sup>1</sup>, Todd M. Morgan<sup>4</sup>, Daniel E. Spratt<sup>5</sup>, Zachery R. Reichert<sup>6</sup>, Rahul Mannan<sup>7</sup>, Xiaoming Wang<sup>7</sup>, Anya Chinnaiyan<sup>7</sup>, Xuhong Cao<sup>7</sup>, Saravana M. Dhanasekaran<sup>7</sup>, Arul M. Chinnaiyan<sup>1,4,7,8,9</sup>, Liron Pantanowitz<sup>1,8</sup> and Rohit Mehra<sup>1,7,8\*</sup> 

## Abstract

**Background:** *TMPRSS2-ERG* gene rearrangement, the most common E26 transformation specific (ETS) gene fusion within prostate cancer, is known to contribute to the pathogenesis of this disease and carries diagnostic annotations for prostate cancer patients clinically. The *ERG* rearrangement status in prostatic adenocarcinoma currently cannot be reliably identified from histologic features on H&E-stained slides alone and hence requires ancillary studies such as immunohistochemistry (IHC), fluorescent in situ hybridization (FISH) or next generation sequencing (NGS) for identification.

**Methods:** Objective: We accordingly sought to develop a deep learning-based algorithm to identify *ERG* rearrangement status in prostatic adenocarcinoma based on digitized slides of H&E morphology alone.

**Design:** *Setting, and Participants:* Whole slide images from 392 in-house and TCGA cases were employed and annotated using QuPath. Image patches of 224 × 224 pixel were exported at 10 ×, 20 ×, and 40 × for input into a deep learning model based on MobileNetV2 convolutional neural network architecture pre-trained on ImageNet. A separate model was trained for each magnification. Training and test datasets consisted of 261 cases and 131 cases, respectively. The output of the model included a prediction of ERG-positive (ERG rearranged) or ERG-negative (ERG not rearranged) status for each input patch.

*Outcome measurements and statistical analysis:* Various accuracy measurements including area under the curve (AUC) of the receiver operating characteristic (ROC) curves were used to evaluate the deep learning model.

**Results and Limitations:** All models showed similar ROC curves with AUC results ranging between 0.82 and 0.85. The sensitivity and specificity of these models were 75.0% and 83.1% (20 × model), respectively.

**Conclusions:** A deep learning-based model can successfully predict *ERG* rearrangement status in the majority of prostatic adenocarcinomas utilizing only H&E-stained digital slides. Such an artificial intelligence-based model can eliminate the need for using extra tumor tissue to perform ancillary studies in order to assess for *ERG* gene rearrangement in prostatic adenocarcinoma.

**Keywords:** Prostate cancer, ERG, Deep learning, Artificial intelligence, Gene fusion, Adenocarcinoma, Whole slide imaging

\*Correspondence: mrohit@med.umich.edu

<sup>1</sup> Department of Pathology, University of Michigan Medical School, Ann Arbor, MI, USA

Full list of author information is available at the end of the article



## Introduction

In medicine today being driven by cutting-edge cancer therapeutics and a personalized approach to delivering healthcare, artificial intelligence (AI) has had an additive impact on the digital transformation in the field. An increasing role of AI, including machine and deep learning methods, is being applied not only to diverse 'omics' fields such as genomics, pharmacogenomics, and proteomics, but also conventional clinical medicine disciplines such as radiology, pathology, immuno-oncology, and others. Substantial published data is accruing indicating that AI can improve diagnoses, offer predictions (e.g. correlation with underlying tumor genomics, theranostic response to treatment), and render prognostic information employing only image data. In the field of genitourinary medicine, deep learning models have been developed to reliably subtype renal cell carcinomas [1] and aid pathologists in the automation of prostate biopsy interpretation [2].

Prostate cancer continues to be a major global cause of morbidity and mortality, with a significant death rate within the United States. The current National Comprehensive Cancer Network (NCCN) guidelines have successfully incorporated genomic tools to guide and improve prostate cancer management. In terms of prostate cancer biology, the discovery of recurrent gene fusions in a majority of prostate cancers has had important clinical and biological implications with a paradigm shift in our understanding of the genomics of common epithelial tumors. Several years ago, our group identified genomic rearrangements in prostate cancer resulting in the fusion of the 5' untranslated end of *TMPRSS2* (Transmembrane serine protease 2; a prostate-specific gene controlled by androgen) to members of the ETS (E26 transformation-specific) family of oncogenic transcription factors, leading to the over-expression of ETS genes like *ERG* (ETS-related gene), *ETV1* (ETS variant transcription factor 1), *ETV4* (ETS variant transcription factor 4), and others, with *ERG* being the most common gene fusion partner [3]. ETS gene fusions are found within a distinct class of prostate cancers that are associated with diagnosis, prognosis, and targeted therapy [4, 5].

Almost half of all prostatic adenocarcinomas, including clinically localized as well as metastatic tumors, are associated with *TMPRSS2-ERG* gene fusion [6]. This causes juxtaposition of the *ERG* gene to androgen-responsive regulatory elements of *TMPRSS2*, that leads to aberrant androgen receptor driven over-expression of ERG protein. Currently, the gold standard for detection of *ERG* gene rearrangement is fluorescent in situ hybridization (FISH) or next generation sequencing (NGS) technology. In routine clinical and surgical pathology practice, while some histologic features associated with prostatic adenocarcinoma such as blue

mucin production and prominent nucleoli in tumor cells have been shown to demonstrate association with underlying *ERG* gene rearrangement [7]. However, these morphologic features are not consistently present in the tumors with *ERG* gene rearrangement and by themselves they are not reliably predictive, hence detection of over-expression of *ERG* protein by immunohistochemistry (IHC) is often used as a surrogate to identify *ERG* gene rearrangement in prostate cancer.

Recent studies have proven largely successful at leveraging AI-based models to recognize and characterize prostate cancer on whole slide imaging (WSI) [8–13]. The training strategies for the AI-models fall under one of two main categories: supervised learning, or weakly supervised learning. The first strategy requires that annotations be made at the pixel-level for each whole slide image. While this approach is meticulous and often arduous for the expert who is annotating the case, it benefits from lower computational burden and fewer overall case requirements for training compared to the latter strategy. However, whenever possible, the weakly supervised deep learning strategy is often preferred due to lower burden on the expert pathologist annotator since this method only requires slide-level annotations i.e. one annotation per slide [14]. The most notable example of a successful weakly supervised deep learning AI-model was developed by Campanella et al. and has gone on to receive the first ever FDA approval for an AI product in Digital Pathology [13].

Morphologic characterization is not the only area in which AI-based models have been successful. Much advancement has been made in predicting genetic mutations based on AI-driven histomorphologic analysis. Deep learning has been used to classify and predict common mutations in a variety of tumors including non-small cell lung carcinomas, bladder urothelial carcinomas, renal cell carcinoma and melanomas. All these methods are based on AI analysis of histology images alone without additional information [15–18]. Given the success of these models, it is feasible that similar deep learning algorithms could be developed to potentially predict underlying genetic aberrations in prostate adenocarcinoma from histopathologic images.

In this study, we accordingly utilized Hematoxylin and Eosin (H&E)-stained whole slide images (WSIs) of prostate adenocarcinoma and sought to develop a deep learning algorithm that could distinguish *ERG* rearranged prostate cancers from those without *ERG* rearrangement. Our results suggest that image features alone can analyze subtle morphological differences between *ERG* gene fusion positive and negative prostate cancers, which would thereby eliminate the need to utilize extra tumor tissue to perform ancillary studies such as IHC, FISH or NGS testing in order to assess for *ERG* gene rearrangement.

## Materials and methods

### Data acquisition and slide scanning

Patient samples were procured from Michigan Medicine and the study was performed under Institutional Review Board-approved protocols. A retrospective pathological and clinical review of radical prostatectomies performed between November 2019 and August 2021 at the University of Michigan Health System was conducted. Only patients without prior history of treatment were included. A total of 163 patients were randomly selected for analysis. Electronic medical records and pathology reports were reviewed to analyze clinical parameters (age at diagnosis, PSA level at diagnosis and treatment modality), and pathological variables (Gleason Score/Grade Group). The H&E-stained glass slides from all cases were re-reviewed by two genitourinary pathologists (VD and RoM) to confirm the diagnosis and evaluate morphologic features. Gleason Score, where applicable, was assigned according to the 2014 modified Gleason grading system and Grade Group was assigned according to the established criteria endorsed by the World Health Organization (WHO). A representative H&E glass slide from each radical prostatectomy was scanned using an Aperio AT2 scanner using 20× objective (Leica Biosystems Inc., Buffalo Grove, IL, USA and 40× magnification (0.25 μm/pixel resolution) was achieved using a 2× optical magnification changer. The scanner generated WSIs as sv5 file format. The microscopic photographs were obtained using an Olympus BX43 microscope with attached camera (Olympus DP47) and cellSens software. ERG rearrangement status for in-house cases was determined by IHC as described below. Out of 163 evaluated cases, 6 cases showed heterogeneity of ERG staining and were hence excluded from further analysis. The remaining 157 in-house cases were included in the final dataset utilized for the purposes of this study.

From the cancer genome atlas (TCGA) database (<https://portal.gdc.cancer.gov>), we downloaded a total of 300 formalin fixed paraffin embedded H&E WSIs of prostate cancer. The images without any discernible cancer morphologically were excluded, and for final analysis 242 images from 235 patients were included. These TCGA

prostate cancer cases were included because they have a known ERG rearrangement status, confirmed from previously reported genomic studies [19].

### Immunohistochemistry

IHC was performed on sections that are selected for scanning for all in-house cases using anti-ERG rabbit monoclonal antibodies (EPR3864, Ventana, prediluted). Appropriate positive and negative controls were included. ERG immunohistochemical expression was performed based on clinically used evaluation criterion where expression of ERG protein within a tumor focus was considered to be positive and such a tumor focus was designated as ERG-positive. Tumor foci which do not expression ERG protein were designated as ERG-negative.

### Deep learning model architecture and evaluation

Regions of tumor from WSIs were manually annotated using QuPath v0.2.3 [20]. The regions annotated as either ERG-positive or ERG-negative were exported as 224 × 224 pixel sized JPEG image patches at 10×, 20×, and 40× magnifications, for input into the deep learning model. For all magnifications image patches were taken from same area. All 235 TCGA cases and 26 in-house cases (n = 261, 67%) were used for training purposes. A separate hold out test dataset, that included the remaining in-house cases (n = 131, 33%), were used for performance evaluation of the model (Table 1). A total of 763,945 patches were generated from regions of interest for the training set and 264,688 patches were generated for the hold-out test set. Patches from the training sets were further randomly subdivided into training, validation, and test subsets with a split ratio of 80:16:4, respectively. Using the Python Keras Application Programming Interface (API), we developed a deep learning algorithm for distinguishing between ERG rearranged and ERG non-rearranged prostate cancer. Development and testing was performed using a computer equipped with an NVIDIA RTX 2070 Super graphics processing unit (GPU) and 16 GB of RAM at 3200 MHz. The algorithm is based on the MobileNetV2 convolutional

**Table 1** Distribution of training and hold-out test datasets utilized for algorithm development

Dataset	Subset	Patients	ERG status		Gleason grade group				
			Positive	Negative	1	2	3	4	5
TCGA cohort	Training set	235	123(52%)	112(48%)	41	68	60	32	34
Internal cohort	Training set	26	11(42%)	15(58%)	1	17	6	0	2
	Hold-out test set	131	60(46%)	71(54%)	0	67	31	4	29

Training subset includes initial training, cross-validation and testing sets. Hold-out test set refers to a separate subset of cases not included as part of the training subset. TCGA The Cancer Genome Atlas

neural network (CNN) architecture pre-trained on ImageNet. The pre-trained MobileNetV2 network was used as the base model. It is preceded by a pre-processing layer which scales input pixel values between -1 and 1 for MobileNetV2. Subsequent to the base model, an additional global average layer, dropout layer, and prediction layer were added. Given the binary nature of this classification task, the prediction layer consisted of a dense layer with a sigmoid activation function. Three different models were trained for each of the different image magnifications (10×, 20×, and 40×). Model weights were fine-tuned using ERG-positive and ERG-negative labeled H&E patches. Data augmentation techniques consisting of horizontal flips, vertical flips, rotations, and contrast variations were applied to the input patches during the training process. After initial training, the hyperparameters were further adjusted using validation set. Finally, models were evaluated using hold-out test set cases, which were independent of training and validation sets, with unlabeled patches as inputs. The output of the models consisted of a prediction of ERG-positive or ERG-negative for each input patches. The workflow used in developing our algorithm is summarized in Fig. 1.

## Results

In this study we developed an AI-based algorithm to establish whether *ERG* gene arrangement can be determined solely from H&E-based histologic images of patients with prostate adenocarcinoma. A total of 26 in-house prostatectomy cases and 235 cases obtained from the TCGA prostate cancer cohort were used for initial training of the model and a separate hold-out test cohort of 131 in-house cases was used for model evaluation (Table 1). The analysis was performed employing 10×, 20× and 40× image magnifications. An overall diagnosis for each WSI was assigned based on percentage of labelled patches favoring an ERG-positive or ERG-negative diagnosis. For each magnification, a cut-off of proportion of ERG positive patches were determined that gives best accuracy. A WSI was labeled as ERG-positive if a proportion of the ERG positive patches were greater than the cut-off determined at a particular magnification. For performance metrics, each WSI was defined as follows: true positive is defined as the correct prediction of ERG-positive cases, false negative as incorrect prediction of ERG-positive cases, true negative as correct prediction of ERG-negative cases, and false positive as incorrect prediction of ERG-negative cases. Representative patches as identified by the algorithm are provided in Fig. 2.

ROC curves were generated for all three models. The best accuracy of 79.4% and an area under the ROC curve of 0.85 were achieved at 20× and 40× magnification, respectively (Fig. 3 and Table 2). A total of 104 out of 131

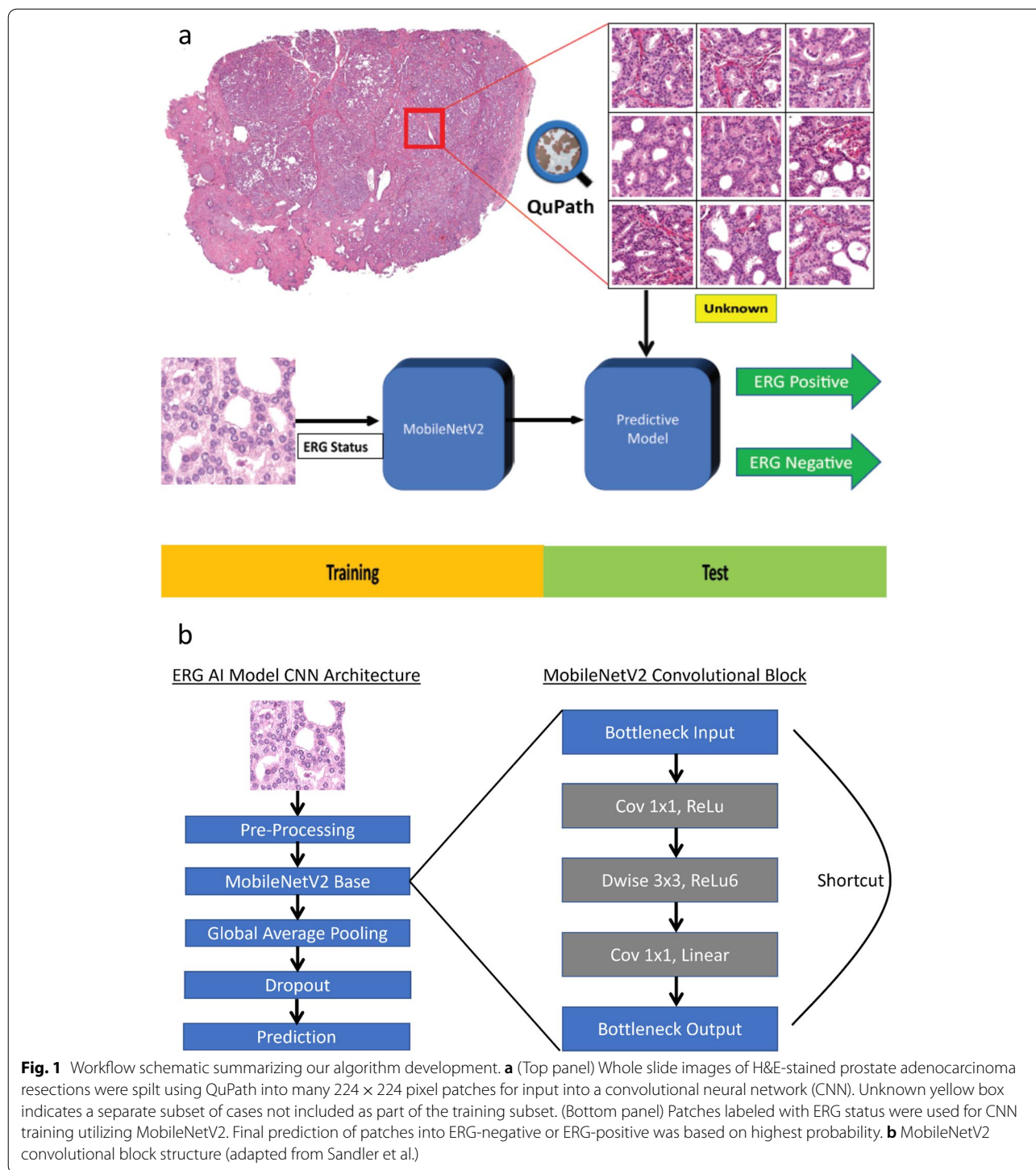
cases (79.4%) were identified correctly by the AI-based algorithm, with a sensitivity of 75.0%, specificity of 83.1%, positive predictive value (PPV) of 78.9%, and negative predictive value (NPV) of 79.7%. The performance of the algorithm at all magnifications was almost equivalent, yielding an accuracy in the range of 78.6% and 79.4% as well as area under the ROC curve in the range of 0.82 and 0.85. The cut-off for ERG-positive patches was 0.5, where best accuracy was achieved for 20× magnification.

The morphologic features of prostatic adenocarcinoma vary greatly according to the grade of the tumor. Hence, we sought to evaluate performance of our model according to different Gleason grades. In order to assess the performance of our model according to different Gleason grades, we evaluated our algorithm separately for subgroups based on their assigned Grade Groups within this cohort. Hold-out cases were categorized into low-grade and high-grade tumors; the lower-grade tumors included Grade Groups 1 and 2, of which these tumors were predominantly comprised of a Gleason grade 3 component; in contrast, the higher-grade tumors were comprised of Grade Groups 3 and higher with the majority of these tumors comprised of Gleason grade 4 and 5 components. With Gleason grades taken into consideration, equivalent accuracy were obtained at all magnifications. For lower-grade tumors the accuracy achieved was 86.6% at 10× and 20× magnifications, while accuracy for higher-grade tumors was 73.4% at 40× magnification. These results are summarized in Table 3.

## Discussion

IHC has become the workhorse of molecular phenotyping for tissues and currently serves as a reliable surrogate to actually performing expensive molecular testing. However, IHC is time-consuming, can be expensive, and dependent on appropriate tissue handling procedures, reagents, and expert laboratory technicians. Furthermore, immunostain findings require visual inspection using a microscope and thus depend on the subjective interpretation of pathologists [21, 22]. Recent technological progress in digital pathology and AI has shown that these new modalities can be used to not only improve efficiency of pathologists, but also provide comparable diagnostic accuracy to pathologists employing traditional light microscopy [12]. Within the domain of surgical pathology, AI-based algorithms can analyze digitized histomorphologic features to effectively distinguish neoplastic and non-neoplastic lesions [23, 24], detect metastasis in lymph nodes [25], predict genomic fusion status within renal neoplasms [26], subtype renal tumors [1], detect prostate cancer in biopsy material [27], as well as grade aggressiveness of certain tumors [4]. To date, AI-based studies have been applied to prostate

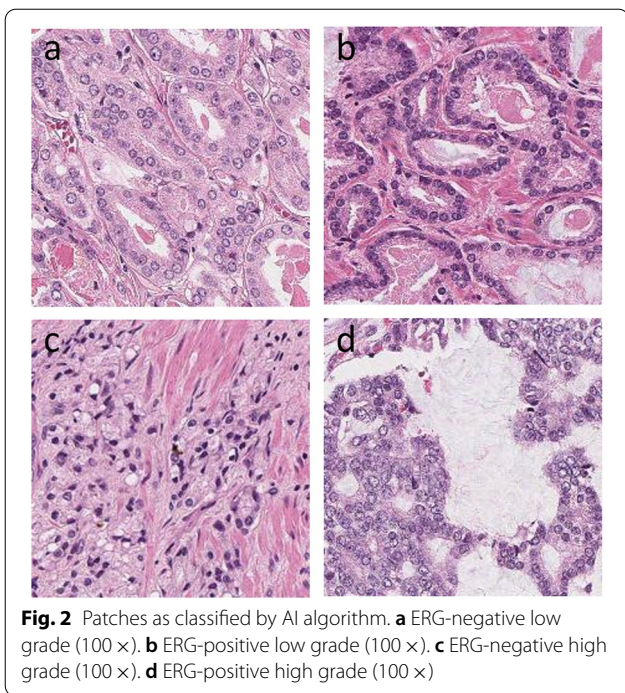




cancer pathology to assist with diagnosis, Gleason grading, prognosis, as well as predict underlying molecular aberrations such as phosphatase and tensin homology (PTEN) loss [2, 28, 29]. In the present study, we developed a deep learning model to predict ERG rearrangement status in patients with prostatic adenocarcinoma.

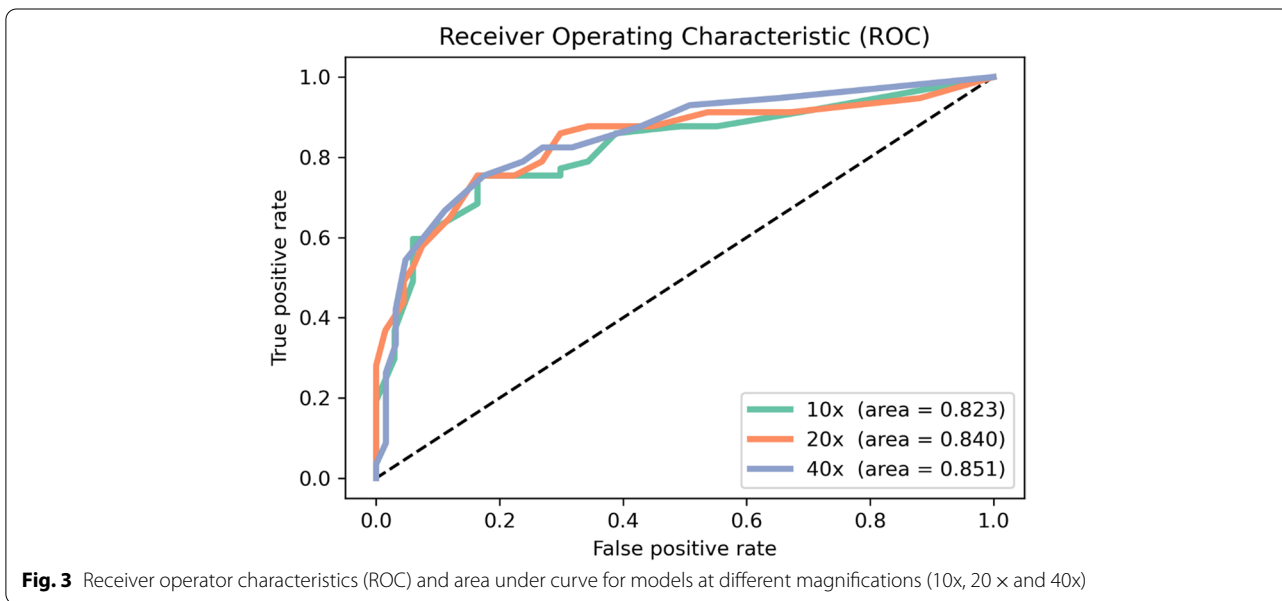
To the best of our knowledge, this is the first study to identify this genomic status directly from scanned H&E-stained slides in surgically resected prostate cancer cases.

*TMPRSS2-ERG* rearrangement, the most common ETS gene fusion in prostate cancer, brings ERG expression under androgen control via androgen receptor-mediated



*TMPRSS2* regulation and results in over-expression of ERG protein [3]. Microscopically, while some *ERG* rearranged prostate cancers are enriched with features such as intraluminal blue mucin and prominent nucleoli, the spectrum of morphology is quite variable and inconsistently predictive of the presence of an *ERG* rearrangement at the genomic level [7]. Hence, it remains challenging to faithfully distinguish prostate cancer with *ERG* rearrangement from those with wild type *ERG* and other molecular subtypes based only on the microscopic evaluation of H&E-stained pathological tissues. As a result, *ERG* gene rearrangement status is usually confirmed by immunohistochemical identification of the overexpression of ERG protein or by dual-color break-apart FISH. However, these ancillary tests require additional time and resources, and they consume precious tissue.

In this study, we demonstrated that a digitized H&E-stained slide analyzed using a deep learning-based model can successfully predict *ERG* fusion status in the majority of prostate cancer cases. We believe that this algorithm can eliminate the need for using extra tumor tissue



**Table 2** Performance metrics of AI-based models at different magnifications

Magnification	AUC	TP	FP	TN	FN	Sensitivity	Specificity	PPV	NPV	Accuracy	F1 score	Cut-off
10 ×	0.82	45	13	58	15	75.0%	81.7%	77.6%	79.5%	78.6%	0.78	0.4
20 ×	0.84	45	12	59	15	75.0%	83.1%	78.9%	79.7%	79.4%	0.79	0.5
40 ×	0.85	45	13	58	15	75.0%	81.7%	77.6%	79.5%	78.6%	0.78	0.35

AUC Area under curve in receiver operator characteristics curve, TP True positive (correctly classified as ERG-positive), FP False positive (incorrectly classified as ERG-positive), TN True negative (correctly classified as ERG-negative), FN False negative (incorrectly classified as ERG-negative), PPV Positive predictive value, NPV Negative predictive value, Cut-off indicates cut-off value of proportion of positive patches that gives best accuracy

**Table 3** Algorithm performance metrics based on tumor grade

Magnification	Grade Group	TP	FP	TN	FN	Sensitivity	Specificity	PPV	NPV	Accuracy	F1 score	Cut-off
10 ×	1–2	33	8	25	1	97.1%	75.8%	80.5%	96.2%	86.6%	0.88	0.4
	3–5	12	5	33	14	46.2%	86.8%	70.6%	70.2%	70.3%	0.56	0.4
20 ×	1–2	28	3	30	6	82.4%	90.9%	90.3%	83.3%	86.6%	0.86	0.5
	3–5	17	9	29	9	65.4%	76.3%	65.4%	76.3%	71.9%	0.65	0.5
40 ×	1–2	28	5	28	6	82.4%	84.8%	84.8%	82.4%	83.6%	0.84	0.35
	3–5	17	8	30	9	65.4%	78.9%	68.0%	76.9%	73.4%	0.67	0.35

TP True positive (correctly classified as ERG-positive), FP False positive (incorrectly classified as ERG-positive), TN True negative (correctly classified as ERG-negative), FN False negative (incorrectly classified as ERG-negative), PPV Positive predictive value, NPV: Negative predictive value, Cut-off indicates cut-off value of proportion of positive patches that gives best accuracy

to perform lengthy and expensive ancillary studies to assess for *ERG* gene rearrangement in prostatic adenocarcinoma. Our AI model was able to accurately predict the presence of an *ERG* gene rearrangement in a large number of cases with varying morphologic patterns and Grade groups, including tumors with low-grade (Grade Group 2 or less) and high-grade features (Grade Group 3 or higher). Higher accuracy was seen in lower-grade tumors. One possibility for this observation may be that higher-grade tumors typically exhibit more diverse morphology.

*ERG* gene rearrangement is known to contribute to the pathogenesis of prostate cancer and provides important clues about the multifocality and metastatic dissemination of this disease. The specificity of this gene rearrangement in prostate cancer allows ERG evaluation by IHC to be of diagnostic value in both primary and metastatic tumors originating from the prostate [6, 30, 31]. *TMPRSS2-ERG* fusions are also prime candidates for the development of new diagnostic assays, including urine-based noninvasive assays [32]. There have been conflicting reports regarding the prognostic value of *ERG* gene rearrangement and its overexpression in prostatic cancer. Hägglöf et al. have demonstrated that high expression of *ERG* is associated with higher Gleason score, aggressive disease and poor survival rates [33]. Similarly, Nam et al. demonstrated that the *TMPRSS2-ERG* fusion gene predicts cancer recurrence after surgical treatment and that this prediction is independent of grade, stage and prostate specific antigen (PSA) levels in blood [34]. Mehra et al. demonstrated *ERG* rearrangement to be associated with a higher stage in prostate cancer [35]. A subsequent study by Fine et al. demonstrated a subset of prostate cancers with *TMPRSS2-ERG* copy number increase, with or without rearrangement, to be associated with higher Gleason score [36]. Nevertheless, other studies have found no association between *TMPRSS2-ERG* fusion and stage, grade, recurrence, or progression [37, 38]. Additionally, in the TCGA dataset, in our limited analyses,

we did not see any other particular genetic mutation that is significantly different between *ERG* rearranged and non-*ERG* rearranged cases. Currently, there is no ERG-targeted therapy approved for treatment of prostate cancer. However, peptidomimetic targeting of transcription factor fusion products has been demonstrated to provide a promising therapeutic strategy for prostate cancer [39]. Previous *ERG* fusion driven biomarker clinical trials utilized interrogation of *ERG* rearrangement status employing IHC or FISH tests [40]. Our study provides a viable and inexpensive alternative to ancillary tissue-based testing methods to detect *ERG* rearrangement status in prostate cancer.

Our study has several strengths and potential limitations. Notable strengths include the use of H&E stained slides only (without the need for concurrent genomic investigation) to predict *ERG* gene fusion status in prostate cancer, utilization of WSI, and employment of diverse datasets including in-house and TCGA datasets with different H&E staining qualities to improve the robustness of our algorithm. This application carries strength in eliminating the need for complex molecular testing utilizing FISH, next-generation sequencing, or molecular surrogate assays like immunohistochemistry; utilizing of H&E slides only allows an easy, economical and efficient methodology to detect ERG gene rearrangement utilizing AI developed model. Importantly, our study paves a foundation for utilizing basic laboratory tools in assessing genomic rearrangements in diverse set of human malignancies (of prostate and other genitourinary tumors).

Computational limitations for both training and test set evaluation were considered when deciding which neural network architecture to utilize. Commonly used architectures for image classification tasks include Inception, VGG16, ResNet50, and MobileNet, among others. Each architecture comes with its own strengths and limitations and each one is designed to be optimal under specific circumstances. For example,

the Inception architecture serves the purpose of reducing computational cost by implementing a shallower network compared to ResNet50 which may negatively impact computational accuracy. MobileNetV2, part of the MobileNet family, further addresses issues of size and speed and is optimally designed for mobile device applications which often require computationally limited platforms. This is accomplished by utilizing 19 inverted residual bottleneck layers following the initial fully convolutional layer, The 19 bottleneck layers are subsequently followed by a point convolutional layer, pooling average layer, and a final convolutional layer. Taking into consideration that these AI-applications are ultimately intended for clinical laboratory settings which may not have access to high-end computational hardware, we ultimately chose a MobileNetV2 architecture pre-trained on ImageNet as our base model due to its balance between accuracy and computational cost [41, 42].

Hardware limitations necessitating relatively small input tiles may contribute to our model's performance. Our training set was relatively enriched in lower-grade tumors as high-grade cancers are less common in daily clinical urological practice. Follow-up studies incorporating more higher-grade tumors will be needed to better assess the performance of our AI-based tool in such scenario. Our algorithm was developed using resection specimens, and further studies would be needed to interrogate findings in biopsy specimens that display smaller volumes of tumor; as a consequence, cut-offs used in this study may need to be adjusted. For the purposes of this study, we did not address disease heterogeneity and multifocality; future studies are likely to address these phenomena.

## Conclusion

We demonstrated that *ERG* rearrangement status in prostate adenocarcinoma can be reliably predicted directly from H&E-stained digital slides utilizing a deep learning algorithm with high accuracy. This approach has great potential to automate digital workflows and avoid using tissue-based ancillary studies to assess for *ERG* gene rearrangement.

## Abbreviations

AI: Artificial intelligence; AUC: Area under curve; ERG: ETS-related gene; ETS: E26 transformation-specific; ETV1: ETS variant transcription factor 1; ETV4: ETS variant transcription factor 4; FISH: Fluorescence in situ hybridization; H&E: Hematoxylin and eosin; IHC: Immunohistochemistry; IRB: Institutional Review Board; mCRPC: Metastatic castrate resistant prostate cancer; NCCN: National Comprehensive Cancer Network; NPV: Negative predictive value; PPV: Positive predictive value; PSA: Prostate specific antigen; PTEN: Phosphatase and tensin homolog; ROC: Receiver operating characteristic; TCGA: The Cancer Genome

Atlas; TMPRSS2: Transmembrane serine protease 2; WHO: World Health Organization; WSI: Whole slide images.

## Acknowledgements

Not applicable.

## Authors' contributions

RoM and LP had full access to all the data in the study and take responsibility for the integrity of the data and accuracy of the data analysis; RoM and LP: Study concept and design; VD, DG, MQY, JC, RaM, XW, AC, XC, and SMD: Acquisition of data; VD, DG, MQY, and RoM: Drafting of the manuscript; TMM, DES, ZRR, and AMC: Critical revision of the manuscript for important intellectual content; VD, MQY, and DG: Statistical analysis; RaM, XW, and CC: Administrative, technical, or material support; RoM and LP: Study supervision; all authors have read and approved the manuscript.

## Funding

No funding was received.

## Availability of data and materials

The TCGA datasets are available in the public domain; rest is not applicable. The code for the model is deposited at <https://doi.org/10.5281/zenodo.5911163>.

## Declarations

### Ethics approval and consent to participate

This study was performed under institutional review board-approved protocols (with a waiver of informed consent). Approval was granted by the Ethics Committee of University of Michigan School of Medicine. (Approval No.: RBMED # 2001–0155).

### Consent for publication

Not applicable.

### Competing interests

LP is on the scientific advisory board for Ibex and NTP and serves as a consultant for Hamamatsu. The University of Michigan has been issued a patent on the detection of ETS gene fusions in prostate cancer, on which RoM and AMC are listed as co-inventors. All other authors have no relevant disclosures.

### Author details

<sup>1</sup>Department of Pathology, University of Michigan Medical School, Ann Arbor, MI, USA. <sup>2</sup>Department of Pathology and Laboratory Medicine, Jackson Memorial Hospital, Miami, FL, USA. <sup>3</sup>Department of Pathology, Vanderbilt University Medical Center, Nashville, TN, USA. <sup>4</sup>Department of Urology, University of Michigan Medical School, Ann Arbor, MI, USA. <sup>5</sup>Department of Radiation Oncology, University Hospitals Seidman Cancer Center, Case Western Reserve University School of Medicine, Cleveland, OH, USA. <sup>6</sup>Department of Medical Oncology, University of Michigan Medical School, Ann Arbor, MI, USA. <sup>7</sup>Michigan Center for Translational Pathology, Ann Arbor, MI, USA. <sup>8</sup>Rogel Cancer Center, Michigan Medicine, Ann Arbor, MI, USA. <sup>9</sup>Howard Hughes Medical Institute, Ann Arbor, MI, USA.

Received: 3 November 2021 Accepted: 17 April 2022

Published online: 05 May 2022

## References

1. M Fenstermaker SA Tomlins K Singh 2020 Development and Validation of a Deep-learning Model to Assist With Renal Cell Carcinoma Histopathologic Interpretation *Urology* 144 152 157 <https://doi.org/10.1016/j.urology.2020.05.094>
2. L Pantanowitz GM Quiroga-Garza L Bien 2020 An artificial intelligence algorithm for prostate cancer diagnosis in whole slide images of core needle biopsies: a blinded clinical validation and deployment study *Lancet Digit Health* 2 e407 e416 [https://doi.org/10.1016/S2589-7500\(20\)30159-X](https://doi.org/10.1016/S2589-7500(20)30159-X)



3. SA Tomlins DR Rhodes S Perner 2005 Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer *Science* 310 644 648 <https://doi.org/10.1126/science.1117679>
4. D Wang DJ Foran J Ren 2015 Exploring Automatic Prostate Histopathology Image Gleason Grading via Local Structure Modeling *Conf Proc IEEE Eng Med Biol Soc* 2015 2649 2652 <https://doi.org/10.1109/EMBC.2015.7318936>
5. AM Udager AM DeMarzo Y Shi 2016 Concurrent nuclear ERG and MYC protein overexpression defines a subset of locally advanced prostate cancer: Potential opportunities for synergistic targeted therapeutics *Prostate* 76 845 853 <https://doi.org/10.1002/pros.23175>
6. R Mehra SA Tomlins J Yu 2008 Characterization of TMPRSS2-ETS gene aberrations in androgen-independent metastatic prostate cancer *Cancer Res* 68 3584 3590 <https://doi.org/10.1158/0008-5472.CAN-07-6154>
7. J-M Mosquera S Perner F Demichelis 2007 Morphological features of TMPRSS2-ERG gene fusion prostate cancer *J Pathol* 212 91 101 <https://doi.org/10.1002/path.2154>
8. Liu Y, An X (2017) A classification model for the prostate cancer based on deep learning. 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). <https://doi.org/10.1109/CISP-BMEI.2017.8302240>
9. E Arvaniti KS Fricker M Moret 2018 Automated Gleason grading of prostate cancer tissue microarrays via deep learning *Sci Rep* 8 12054 <https://doi.org/10.1038/s41598-018-30535-1>
10. Nagpal K, Foote D, Liu Y, et al (2019) Development and validation of a deep learning algorithm for improving Gleason scoring of prostate cancer. *npj Digit Med* 2:1–10. <https://doi.org/10.1038/s41746-019-0112-2>
11. Tolkach Y, Dohmgörger T, Toma M, Kristiansen G (2020) High-accuracy prostate cancer pathology using deep learning. *Nature Machine Intelligence* 2:411–418. <https://doi.org/10.1038/s42256-020-0200-7>
12. G Campanella MG Hanna L Geneslaw 2019 Clinical-grade computational pathology using weakly supervised deep learning on whole slide images *Nat Med* 25 1301 1309 <https://doi.org/10.1038/s41591-019-0508-1>
13. Mun Y, Paik I, Shin S-J, et al (2021) Yet Another Automated Gleason Grading System (YAAGGS) by weakly supervised deep learning. *npj Digit Med* 4:1–9. <https://doi.org/10.1038/s41746-021-00469-6>
14. J Laak van der F Ciompi G Litjens 2019 No pixel-level annotations needed *Nat Biomed Eng* 3 855 856 <https://doi.org/10.1038/s41551-019-0472-6>
15. N Coudray PS Ocampo T Sakellaropoulos 2018 Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning *Nat Med* 24 1559 1567 <https://doi.org/10.1038/s41591-018-0177-5>
16. N Coudray A Tsirigos 2020 Deep learning links histology, molecular signatures and prognosis in cancer *Nat Cancer* 1 755 757 <https://doi.org/10.1038/s43018-020-0099-2>
17. CS Velmahos M Badgeley Y-C Lo 2021 Using deep learning to identify bladder cancers with FGFR-activating mutations from histology images *Cancer Med* 10 4805 4813 <https://doi.org/10.1002/cam4.4044>
18. Kim RH, Nomikou S, Dawood Z, et al (2019) A Deep Learning Approach for Rapid Mutational Screening in Melanoma. *bioRxiv* 610311. <https://doi.org/10.1101/610311>
19. Cancer Genome Atlas Research Network 2015 The Molecular Taxonomy of Primary Prostate Cancer *Cell* 163 1011 1025 <https://doi.org/10.1016/j.cell.2015.10.025>
20. P Bankhead MB Loughrey JA Fernández 2017 QuPath: Open source software for digital pathology image analysis *Sci Rep* 7 16878 <https://doi.org/10.1038/s41598-017-17204-5>
21. KB Engel HM Moore 2011 Effects of preanalytical variables on the detection of proteins by immunohistochemistry in formalin-fixed, paraffin-embedded tissue *Arch Pathol Lab Med* 135 537 543 <https://doi.org/10.1043/2010-0702-RAIR.1>
22. JA Robb L Bry PM Sluss 2015 A Call to Standardize Preanalytic Data Elements for Biospecimens, Part II *Arch Pathol Lab Med* 139 1125 1128 <https://doi.org/10.5858/arpa.2014-0572-CP>
23. G Litjens CI Sánchez N Timofeeva 2016 Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis *Sci Rep* 6 26286 <https://doi.org/10.1038/srep26286>
24. D Romo-Bucheli A Janowczyk H Gilmore 2017 A deep learning based strategy for identifying and associating mitotic activity with gene expression derived risk categories in estrogen receptor positive breast cancers *Cytometry A* 91 566 573 <https://doi.org/10.1002/cyto.a.23065>
25. AN Basavanthally S Ganesan S Agner 2010 Computerized image-based detection and grading of lymphocytic infiltration in HER2+ breast cancer histopathology *IEEE Trans Biomed Eng* 57 642 653 <https://doi.org/10.1109/TBME.2009.2035305>
26. Cheng J, Han Z, Mehra R, et al (2020) Computational analysis of pathological images enables a better diagnosis of TFE3 Xp11.2 translocation renal cell carcinoma. *Nat Commun* 11:1778. <https://doi.org/10.1038/s41467-020-15671-5>
27. P Raciti J Sue R Ceballos 2020 Novel artificial intelligence system increases the detection of prostate cancer in whole slide images of core needle biopsies *Mod Pathol* 33 2058 2066 <https://doi.org/10.1038/s41379-020-0551-y>
28. Leo P, Chandramouli S, Farré X, et al (2021) Computationally Derived Cribriform Area Index from Prostate Cancer Hematoxylin and Eosin Images Is Associated with Biochemical Recurrence Following Radical Prostatectomy and Is Most Prognostic in Gleason Grade Group 2. *Eur Urol Focus* S2405–4569(21)00122-X. <https://doi.org/10.1016/j.euf.2021.04.016>
29. SA Harmon PG Patel TH Sanford 2021 High throughput assessment of biomarkers in tissue microarrays using artificial intelligence: PTEN loss as a proof-of-principle in multi-center prostate cancer cohorts *Mod Pathol* 34 478 489 <https://doi.org/10.1038/s41379-020-00674-w>
30. J Fontugne K Davis N Palanisamy 2016 Clonal evaluation of prostate cancer foci in biopsies with discontinuous tumor involvement by dual ERG/SPINK1 immunohistochemistry *Mod Pathol* 29 157 165 <https://doi.org/10.1038/modpathol.2015.148>
31. R Mehra B Han SA Tomlins 2007 Heterogeneity of TMPRSS2 gene rearrangements in multifocal prostate adenocarcinoma: molecular evidence for an independent group of diseases *Cancer Res* 67 7991 7995 <https://doi.org/10.1158/0008-5472.CAN-07-2043>
32. SA Tomlins JR Day RJ Lonigro 2016 Urine TMPRSS2:ERG Plus PCA3 for Individualized Prostate Cancer Risk Assessment *Eur Urol* 70 45 53 <https://doi.org/10.1016/j.eururo.2015.04.039>
33. C Hägglöf P Hammarsten K Strömvall 2014 TMPRSS2-ERG Expression Predicts Prostate Cancer Survival and Associates with Stromal Biomarkers *PLoS ONE* 9 e86824 <https://doi.org/10.1371/journal.pone.0086824>
34. RK Nam L Sugar W Yang 2007 Expression of the TMPRSS2:ERG fusion gene predicts cancer recurrence after surgery for localised prostate cancer *Br J Cancer* 97 1690 1695 <https://doi.org/10.1038/sj.bjc.6604054>
35. R Mehra SA Tomlins R Shen 2007 Comprehensive assessment of TMPRSS2 and ETS family gene aberrations in clinically localized prostate cancer *Mod Pathol* 20 538 544 <https://doi.org/10.1038/modpathol.3800769>
36. SW Fine A Gopalan MA Leversha 2010 TMPRSS2-ERG gene fusion is associated with low Gleason scores and not with high-grade morphological features *Mod Pathol* 23 1325 1333 <https://doi.org/10.1038/modpathol.2010.120>
37. A Gopalan MA Leversha JM Satagopan 2009 TMPRSS2-ERG gene fusion is not associated with outcome in patients treated by prostatectomy *Cancer Res* 69 1400 1406 <https://doi.org/10.1158/0008-5472.CAN-08-2467>
38. LM FitzGerald I Agalliu K Johnson 2008 Association of TMPRSS2-ERG gene fusion with clinical characteristics and outcomes: results from a population-based study of prostate cancer *BMC Cancer* 8 230 <https://doi.org/10.1186/1471-2407-8-230>
39. X Wang Y Qiao IA Asangani 2017 Development of Peptidomimetic Inhibitors of the ERG Gene Fusion Product in Prostate Cancer *Cancer Cell* 31 532 548.e7 <https://doi.org/10.1016/j.ccell.2017.02.017>
40. M Hussain S Daignault-Newton PW Twardowski 2018 Targeting Androgen Receptor and DNA Repair in Metastatic Castration-Resistant Prostate Cancer: Results From NCI 9012 *J Clin Oncol* 36 991 999 <https://doi.org/10.1200/JCO.2017.75.7310>
41. Sandler M, Howard A, Zhu M, et al (2018) MobileNetV2: Inverted Residuals and Linear Bottlenecks. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, Salt Lake City, UT, pp 4510–4520 <https://doi.org/10.48550/arXiv.1801.04381>
42. Deng J, Dong W, Socher R, et al (2009) ImageNet: A large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, Miami, FL, pp 248–255 <https://doi.org/10.1109/CVPR.2009.5206848>

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.