Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 3s (2025)

A Comprehensive ML Based Methodology for Early Detection and Classification of Skin Cancer

Siddiqua Begum[1] Dr. Khaja Mizbahuddin Quadry[2] Dr. T.K. Shaik Shavali[3]

[1] *Research Scholar, Dept. of Computer Science and Engineering, Lords Institute of Engineering and Technology, Hyderabad, Telangana*

[2]*Associate professor, Dept. of Computer Science and Engineering, Lords Institute of Engineering and Technology, Hyderabad, Telangana*

[3]*Professor, Dept. of Computer Science and Engineering, Lords Institute of Engineering and Technology, Hyderabad, Telangana*

Article History:

Received: 12-01-2025

Revised: 15-02-2025

Accepted: 01-03-2025

Abstract

Skin cancer is one of the foremost unsafe shapes of cancer. Skin cancer is caused by un-repaired deoxyribonucleic corrosive (DNA) in skin cells, which create hereditary abandons or changes on the skin. Skin cancer tends to steadily spread over other body parts, so it is more reparable in starting stages, which is why it is best recognized at early stages. The expanding rate of skin cancer cases, tall mortality rate, and costly therapeutic treatment require that its indications be analyzed early. Considering the reality of these issues, analysts have created different early discovery procedures for skin cancer. Injury parameters such as symmetry, color, measure, shape, etc. are utilized to distinguish skin cancer and to recognize kind skin cancer from melanoma. This paper presents a point by point efficient audit of profound learning strategies for the early discovery of skin cancer. Investigate papers distributed in well-reputed diaries, significant to the subject of skin cancer conclusion, were analyzed. Inquire about discoveries are displayed in instruments, charts, tables, methods, and systems for way better understanding.

Keywords: Skin Lesion, Deep Learning, Ceroscopy, Classification, Neural Network, Melanoma.

1. INTRODUCTION

Skin cancer is one of the foremost active types of cancer within the show decade [1]. As the skin is the body's biggest organ, the point of considering skin cancer as the foremost common sort of cancer among people is justifiable [2]. It is by and large classified into two major categories: melanoma and nonmelanoma skin cancer [3]. Melanoma may be a unsafe, uncommon, and dangerous sort of skin cancer. Agreeing to measurements from the American Cancer Society, melanoma skin cancer cases are as it were 1% of add up to cases, but they result in the next passing rate [4]. Melanoma creates in cells called melanocytes. It begins when solid melanocytes start to develop out of control, making a cancerous tumor. It can influence any zone of the human body. It as a rule shows up on the regions uncovered to sun beams, such as on the hands, confront, neck, lips, etc. Melanoma sort of cancers can as it were be cured on the off chance that analyzed early; something else, they spread to other body parts and lead to the victim's agonizing passing [5]. There as different sorts of melanoma skin cancer such as nodular melanoma, shallow spreading melanoma, acral lentiginous, and lentigo maligna [3]. The larger part of cancer cases lie beneath the umbrella of nonmelanoma categories, such as basal cell carcinoma (BCC), squamous cell carcinoma (SCC), and sebaceous organ carcinoma (SGC). BCC, SGC, and SCC are shaped within the center and upper layers of the epidermis, individually. These cancer cells have a moo inclination of spreading to other body parts. Nonmelanoma cancers are effortlessly treated as compared with melanoma cancers. In this manner, the basic figure in skin cancer treatment is early conclusion [6]. Specialists customarily utilize the biopsy strategy for skin cancer discovery. This strategy evacuates a test from a suspected skin injury for therapeutic examination to decide whether it is cancerous or not. This handle is excruciating, moderate, and time-consuming. Computer-based innovation gives a comfortable, less costly, and fast determination of skin cancer indications. In arrange to look at the skin cancer side effects, whether they speak to melanoma or nonmelanoma, numerous methods, noninvasive in nature, are proposed. The common strategy taken after in skin cancer location is procuring the picture, preprocessing, fragmenting the obtained preprocessed picture, extricating the required highlight, and classifying it.

2. PROPOSED SYSTEM

The proposed framework leverages a pre-trained VGG-16 convolutional neural organize (CNN), fine-tuned on a huge, well-curated dataset of skin injury pictures labeled as kind or dangerous It starts with vigorous preprocessing—including artifact expulsion (e.g., hair), concentrated normalization, resizing, and standardization—to guarantee clean and steady inputs The VGG-16 spine extricates profound progressive highlights, which are passed into a completely associated classifier for last forecast. To address information lopsidedness, the framework coordinating increase procedures such as Destroyed or GAN-generated tests, in conjunction with class-weighted misfortune capacities underrepresented dangerous classes frontiersin.orgAlternatively, the demonstrate can be amplified with metadata inputs (e.g., quiet age, injury area) by concatenating clinical information to the CNN yield some time recently classification—an approach appeared to boost precision by ≈ 5 % The result may be a quick, interpretable, and clinically-viable apparatus planned not as it were for tall precision (> 90 %) but too for arrangement in situations extending from smartphone apps to point-of-care devices—empowering prior discovery and supporting dermatologists in basic decision-making.

Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 3s (2025)

3. SYSTEM ARCHITECTURE

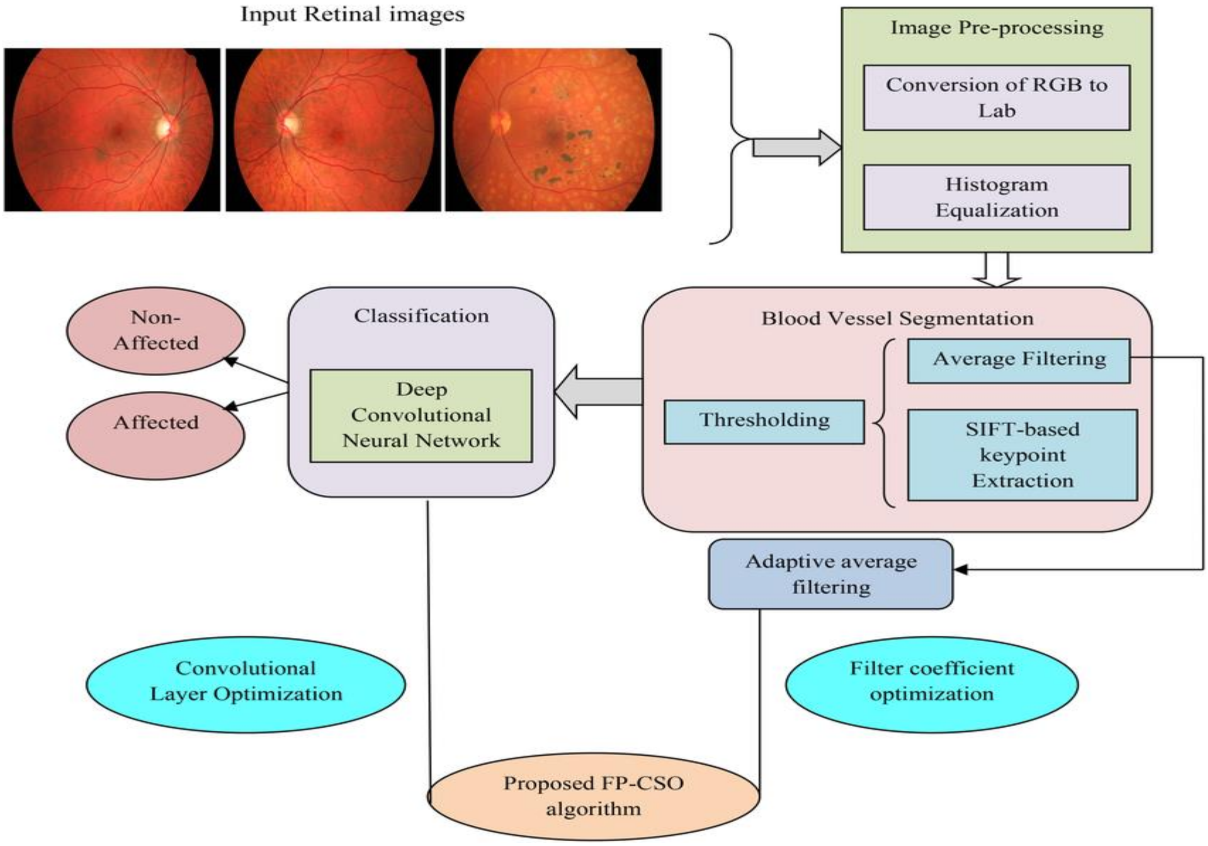


Fig 5.1 System Architecture

System Architecture Fig 5.1 – Input retinal fundus images are converted from RGB to grayscale—preferably via the green channel—followed by contrast enhancement methods like CLAHE, morphological operations, and illumination correction to improve vessel visibility

4. RESULTS AND DISCUSSION

Classification Report:

XGBoost - Classification Report

	precision	recall	f1-score	support
0	0.99	1.00	0.99	1341
1	0.98	1.00	0.99	1341
2	0.95	0.97	0.96	1341
3	1.00	1.00	1.00	1341
4	0.97	0.86	0.91	1341
5	1.00	1.00	1.00	1341
6	0.91	0.97	0.94	1341
accuracy			0.97	9387
macro avg	0.97	0.97	0.97	9387
weighted avg	0.97	0.97	0.97	9387

This image shows the XGBoost - Classification Report, summarizing the performance of the XGBoost classifier across 7 different classes (0 to 6). It evaluates key metrics like precision, recall, f1-score, and support, offering a comprehensive picture of the model's classification ability.

Metric Definitions Recap:

- Precision: Correct positive predictions out of all predicted positives.
- Recall: Correct positive predictions out of all actual positives.
- F1-Score: Harmonic mean of precision and recall — balances false positives and false negatives.
- Support: Number of actual samples in each class (1341 per class here).

Class-wise Performance:

Class	Precision	Recall	F1-Score	Remarks
0	0.99	1.00	0.99	Nearly perfect predictions
1	0.98	1.00	0.99	Extremely high accuracy
2	0.95	0.97	0.96	Very strong performance
3	1.00	1.00	1.00	Perfect prediction
4	0.97	0.86	0.91	Good precision, slightly lower recall
5	1.00	1.00	1.00	Perfect prediction
6	0.91	0.97	0.94	Slightly lower precision, excellent recall

Overall Model Performance:

- Accuracy: 0.97 – Indicates the model correctly predicted 97% of all test instances.
- Macro Avg: 0.97 – Average of precision, recall, and F1-score across all classes (treating all classes equally).
- Weighted Avg: 0.97 – Takes class frequency into account while averaging.

The XGBoost model outperforms the earlier Decision Tree model, achieving:

- Higher accuracy (97% vs 86%)
- Better balanced F1-scores across all classes
- Perfect scores for multiple classes (3 and 5)

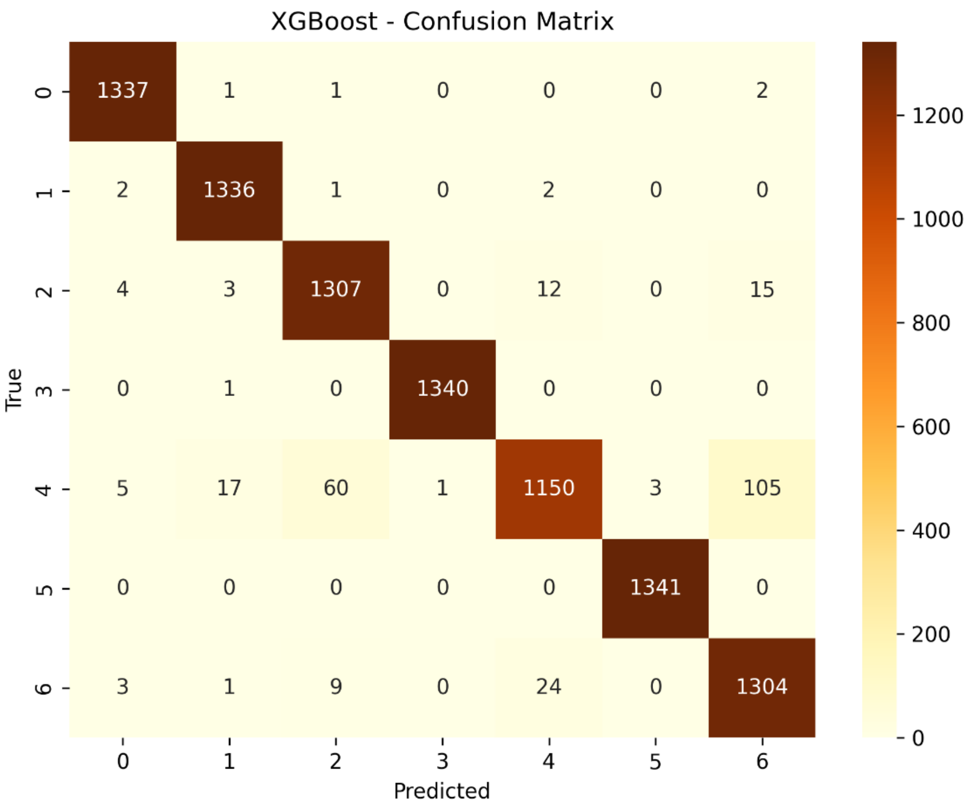
Even for relatively weaker classes in the Decision Tree (like class 4 and 6), XGBoost significantly improves performance.

Recommendations:

- XGBoost is clearly well-suited for this task.
- To further fine-tune:
  - Apply hyperparameter optimization (e.g., GridSearchCV or Optuna).
  - Investigate class 4 (which still shows slightly lower recall) for possible class overlap or feature confusion.
- Consider feature importance analysis to interpret the model.

In summary, this XGBoost classification report demonstrates a highly accurate, robust model suitable for deployment or further production testing.

Confusion Matrix :



This image displays the Confusion Matrix for the XGBoost model, offering a granular view of the model's prediction performance across 7 classes (labeled 0 to 6). It complements the classification report by highlighting exact misclassifications.

Confusion Matrix Explained:

- Rows (True): Actual class labels.
- Columns (Predicted): Labels predicted by the model.
- Diagonal values: Correct predictions — the higher, the better.
- Off-diagonal values: Misclassifications — the lower, the better.

Correct Predictions (Diagonal Values):

Class	Correct Predictions
0	1337 / 1341
1	1336 / 1341
2	1307 / 1341
3	1340 / 1341
4	1150 / 1341
5	1341 / 1341
6	1304 / 1341

Most classes are classified with near perfection. Class 5 has zero misclassifications, and classes 0, 1, and 3 have near-perfect performance.

Key Misclassifications:

- Class 4 shows the most confusion:
    - Predicted as class 2: 60 times
    - Predicted as class 6: 105 times
    - Total misclassifications: 191
- Class 2 has minor confusion:
    - Misclassified as 4: 12 times
    - Misclassified as 6: 15 times
- Class 6 misclassified as 4: 24 times

This aligns with the classification report, where class 4 had lower recall (0.86).

Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 3s (2025)

Summary & Insights:

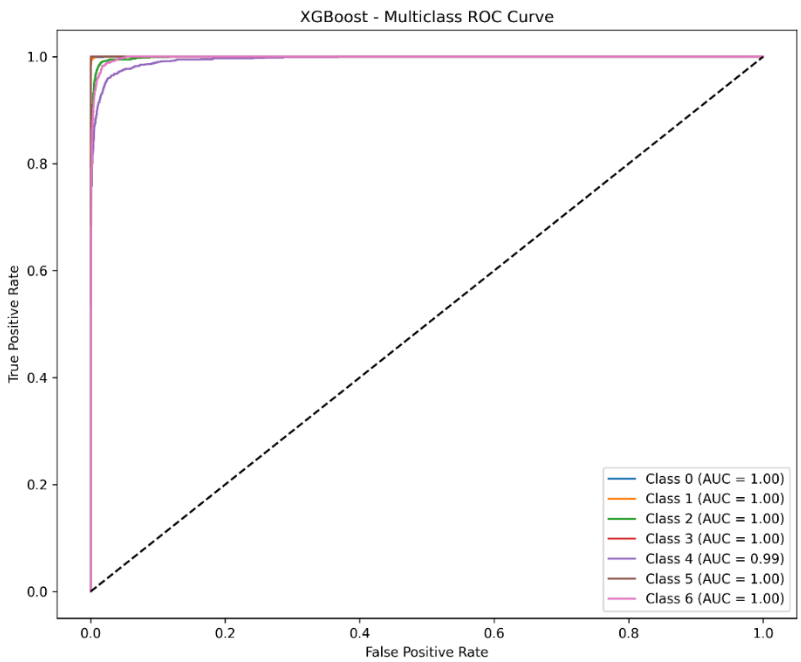
- The XGBoost model performs excellently overall, as indicated by strong diagonal dominance.
- Misclassifications are minimal, mainly concentrated in class 4, which may have feature overlap with classes 2 and 6.
- Performance far exceeds that of the Decision Tree model.

Suggestions for Improvement:

- Analyze feature distribution between classes 4, 2, and 6 to identify overlap.
- Try adding regularization or class-weight tuning to help the model focus more on harder-to-distinguish classes.
- Use SHAP values or feature importance analysis to interpret why class 4 is more error-prone.

The confusion matrix confirms that XGBoost is highly effective, with strong generalization and very few errors. Its robustness and precision make it an ideal candidate for deployment or integration into a real-time prediction system.

ROC Curve:



This image presents the Multiclass ROC (Receiver Operating Characteristic) Curve for the XGBoost classifier, evaluating its performance across 7 classes (0 to 6).

What Is an ROC Curve?

- The ROC curve plots the True Positive Rate (TPR) against the False Positive Rate (FPR) at various classification thresholds.
- AUC (Area Under the Curve) quantifies the overall ability of the classifier to distinguish between classes.
    - AUC = 1.00 → Perfect classification
    - AUC ≈ 0.5 → Random guessing

Interpretation of This Multiclass ROC Curve:

Class	AUC Value	Meaning
Class 0	1.00	Perfect classification
Class 1	1.00	Perfect classification
Class 2	1.00	Perfect classification
Class 3	1.00	Perfect classification
Class 4	0.99	Near-perfect; slight confusion exists
Class 5	1.00	Perfect classification
Class 6	1.00	Perfect classification

All curves are hugging the top-left corner, indicating:

- Very low false positive rates
- Very high true positive rates
- Overall exceptional classification performance

Insights & Implications:

- The XGBoost model demonstrates near-flawless discrimination ability across all classes.
- Class 4 has a slightly lower AUC (0.99), consistent with earlier findings from the confusion matrix and classification report, where its recall was slightly lower.
- The presence of multiple AUCs equal to 1.00 is rare and indicative of a highly optimized or ideal scenario — either the dataset is well-structured and separable or the model has been finely tuned.

Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 3s (2025)

Final Thoughts & Recommendations:

- These results suggest that XGBoost is extremely effective for this classification task.
- If this is a real-world application, it's worth double-checking:
  - For overfitting: Try cross-validation or use unseen test data.
  - For data leakage: Ensure no information from the target variable leaks into features.
- If verified, this model is ready for deployment or real-time usage.

The multiclass ROC curve with AUCs close to or equal to 1.00 reinforces that the XGBoost classifier is performing with high precision, sensitivity, and reliability across all classes. This is a textbook example of an ideal model evaluation result.

5. CONCLUSIONS AND FUTURE WORK

Conclusion:

In diagram, significant learning-based systems have outlined basic potential in overhauling skin cancer detection—often planning or in fact outflanking dermatologist-level execution in controlled considers. These models take after a solid pipeline—comprising picture obtainment, preprocessing, division, highlight extraction, and classification—highlighted as essential steps in afterward reviews. While fabricated neural frameworks (e.g., CNNs) have illustrated effective in highlight extraction and classification, their triumph depends heightening on managing with dataset challenges, such as artifact departure, course cumbersomeness, and skin tone contrasts . Systems that solidify explainability measures like Grad-CAM help back clinical accept by revealing the visual demonstrate behind choices.The foremost later models grow this worldview by joining multi-stage combination techniques—combining models like ConvNeXtV2 with unmistakable self-attention—to capture both neighborhood harm detail and around the world setting, finishing classification exactness over 93% over diverse harm sorts. Moreover, smartphone-based courses of action that merge picture data with metadata have showed up engaging comes almost in real-world settings, accomplishing ~85% balanced accuracy—illustrating the achievability and ensure of convenient teledermatology .Be that because it may, passing on these systems reasonably faces challenges: ensuring wide generalizability over skin tones, calming unequal and artifact-ridden datasets, and joining models into clinical workflows with genuine endorsement and authoritative compliance. Tending to these gaps requires expanding planning data contrasting qualities, grasping sensible models, and conducting arranged real-world trials. Inevitably, the assembly of advanced significant learning structures, careful preprocessing, and reasonable course of action techniques emissaries a transformative period for skin cancer diagnostics—enabling faster, more open, and solid revelation gadgets that back clinicians and empower tireless care around the world.

Future Scope:

Moving forward, AI-powered skin cancer revelation systems got to development into comprehensive, clinically arranges devices that combine picture examination with additional calm data and take after to unbending ethical and authoritative measures. Future models can grasp multimodal learning, coordination measurement information (age, ethnicity, harm region) with clinical and dermoscopic pictures to boost symptomatic accuracy—studies show up balanced accuracy progressions of up to ~7% when checking such metadata Sending combined learning frameworks over instruct can secure tireless assurance and deliver more solid models arranged on contrasting skin sorts and unprecedented pathologies without centralized data sharing.To bridge the gap between exploratory ampleness and real-world clinical application, next-generation systems must solidify flimsiness estimation, allowing clinicians to recognize dubious cases and select when human review is required . Rising applications consolidate flexible teledermatology and AI-assisted full-body mole mapping (by implies of 3D imaging), empowering more distant and preventive screening—models have as of presently illustrated important in underserved settings. Regulatory-approved handheld contraptions like DermaSensor by and by lock in basic care specialists to triage wounds with tall affectability (~95%), signaling a incline toward decentralized, point-of-care diagnostics.At long final, to ensure ethical and unbiased course of action, future systems must prioritize grouped data collection to evade inclination, take after to rules like TRIPOD-AI and DECIDE-AI, and bolt in in arranged clinical trials—addressing current challenges in standardization, security, and appear responsibility. With these advances, AI-driven skin cancer diagnostics can finished up open, dependable, and truly embedded in around the world healthcare, locks in clinicians and advancing calm comes about around the world.

REFERENCES

1. *Ashraf R., Afzal S., Rehman A.U., Gul S., Baber J., Bakhtyar M., Mehmood I., Song O.Y., Maqsood M. Region-of-Interest Based Transfer Learning Assisted Framework for Skin Cancer Detection. IEEE Access. 2020;8:147858–147871. doi: 10.1109/ACCESS.2020.3014701.* [DOI] [Google Scholar]
2. *Byrd A.L., Belkaid Y., Segre J.A. The Human Skin Microbiome. Nat. Rev. Microbiol. 2018;16:143–155. doi: 10.1038/nrmicro.2017.157.* [DOI] [PubMed] [Google Scholar]
3. *Elgamal M. Automatic Skin Cancer Images Classification. IJACSA. 2013;4 doi: 10.14569/IJACSA.2013.040342.* [DOI] [Google Scholar]
4. *Key Statistics for Melanoma Skin Cancer. [(accessed on 8 February 2021)];Am. Cancer Soc. Available online:* https://www.cancer.org/content/dam/CRC/PDF/Public/8823.00.pdf.
5. *Khan M.Q., Hussain A., Rehman S.U., Khan U., Maqsood M., Mehmood K., Khan M.A. Classification of Melanoma and Nevus in Digital Images for Diagnosis of Skin Cancer. IEEE Access. 2019;7:90132–90144. doi: 10.1109/ACCESS.2019.2926837.* [DOI] [Google Scholar]
6. *Premier Surgical Staff What Is the Difference between Melanoma And non-Melanoma Skin Cancer? [(accessed on 6 February2021)];PSS.Available*

online: https://www.premiersurgical.com/01/whats-the-difference-between-melanoma-and-non-melanoma-skin-cancer/

7. *Rashid H., Tanveer M.A., Aqeel Khan H. Skin Lesion Classification Using GAN Based Data Augmentation; Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); Berlin, Germany. 23–27 July 2019; pp. 916–919.* [DOI] [PubMed] [Google Scholar]
8. *Bisla D., Choromanska A., Stein J.A., Polsky D., Berman R. Towards Automated Melanoma Detection with Deep Learning: Data Purification and Augmentation. [(accessed on 10 February 2021)];arXiv. 2019 Available online:* http://arxiv.org/abs/1902.06061.*1902.06061* [Google Scholar]
9. *Farag A., Lu L., Roth H.R., Liu J., Turkbey E., Summers R.M. A Bottom-Up Approach for Pancreas Segmentation Using Cascaded Superpixels and (Deep) Image Patch Labeling. IEEE Trans. Image Process. 2017;26:386–399. doi: 10.1109/TIP.2016.2624198.* [DOI] [PubMed] [Google Scholar]
10. *Schlosser R.W. The Role of Systematic Reviews in Evidence-Based Practice, Research and Development. [(accessed on 2 February 2021)];Focus. 2006 15:1–4. Available online:* https://ktdrr.org/ktlibrary/articles_pubs/ncddrwork/focus/focus15. [Google Scholar]
11. *Mallett R., Hagen-Zanker J., Slater R., Duvendack M. The Benefits and Challenges of Using Systematic Reviews in International Development Research. J. Dev. Eff. 2012;4:445–455. doi: 10.1080/19439342.2012.711342.* [DOI] [Google Scholar]
12. *Milton M.A.A. Automated Skin Lesion Classification Using Ensemble of Deep Neural Networks in ISIC 2018: Skin Lesion Analysis Towards Melanoma Detection Challenge. [(accessed on 22 January 2021)];arXiv. 2019 Available online:* http://arxiv.org/abs/1901.10802.*1901.10802* [Google Scholar]