Wadi Alshatti University Journal of
Pure and Applied Sciences

مجلة جامعة وادي الشاطئ
للعلوم البحتة والتطبيقية

Volume 3, No. 2, July-December 2025 Online ISSN: 3006-0877 المجلد 3، الاصدار 2، يوليو- ديسمبر 2025

RESEARCH ARTICLE **DIGITAL IMAGE PROCESSING**

Multi-Class Classification of Skin Cancer Images Using a Deep Learning-Based Convolutional Neural Network (CNN)

Fatima Abdulnabi Ahmed[1*], Al-Zahraa Hassan Othman[1], Ali Abdrhman Ukasha[1]

[1]Department of Electrical and Electronic Engineering, Faculty of Engineering, Wadi Alshatti University, Brack, Libya

ARTICLE HISTORY
Received 31 May 2025
Revised 30 August 2025
Accepted 02 September 2025
Online 04 September 2025

KEYWORDS
Convolutional neural network;
Deep leaning;
Multi-class classification;
Random oversampling;
Skin cancer.

ABSTRACT

Skin cancer is one of the most common and serious types of cancer, often resulting from the transformation of melanocytes due to excessive exposure to ultraviolet radiation. Early diagnosis significantly improves survival rates and reduces mortality. However, traditional visual diagnosis faces challenges in distinguishing between the seven main categories of skin conditions due to visual similarities and data imbalance. This study aims to develop an automated diagnostic model based on a Convolutional Neural Network (CNN) to enhance classification accuracy and reduce reliance on human evaluation. The HAM10000 dataset (Human Against Machine with 10,000 Training Images), comprising 10,015 images across seven categories, was used, with class imbalance addressed through Random Oversampling. The data were split into 80% for training and 20% for testing, with 20% of the training set reserved for validation. The model was constructed using sequential Conv2D and MaxPooling layers for feature extraction, followed by Flatten and Dense layers for final classification. Training was performed with the Adam optimizer for adaptive weight adjustment and the Categorical Crossentropy loss function suitable for multi-class classification. The model achieved an overall accuracy of 95.45% and demonstrated strong performance across all categories based on Precision, Recall, and F1-Score, with Precision ranging from 84% to 99% and a weighted average of 96% for both Precision and F1-Score. These results highlight the proposed model's clinical significance as an intelligent decision-support tool, enhancing early diagnosis, reducing human error, and improving healthcare quality.

تصنيف متعدد الفئات لصورسرطان الجلد باستخدام الشبكة العصبية التلافيفية (CNN) القائمة على التعلم العميق

فاطمة عبد النبي أحمد [1*]، الزهراء حسن عثمان [1]، علي عبد الرحمن عكاشة [1]

الكلمات المفتاحية
الشبكة العصبية التلافيفية
التعلم العميق
التصنيف متعدد الفئات
التجاوز العشوائي للعينات
سرطان الجلد

الملخص

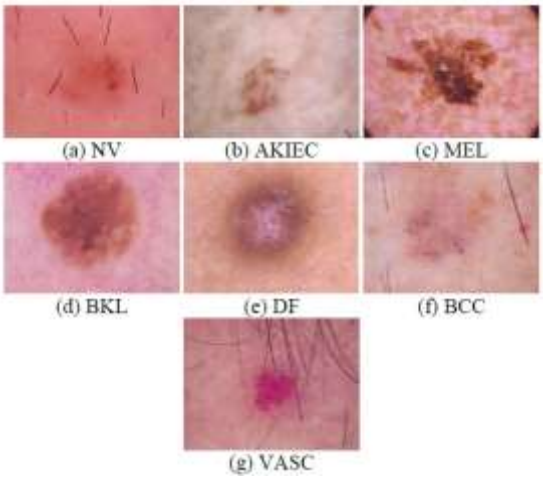
يُعَدّ سرطان الجلد من أكثر أنواع السرطانات انتشارًا وخطورةً، وينجم غالبًا عن تحوّل الخلايا الصبغية إثر التعرّض المفرط للأشعة فوق البنفسجية، ويُسهم التشخيص المبكر في رفع معدلات الشفاء بشكل ملحوظ وتقليل نسب الوفيات المرتبطة به . ومع ذلك، يواجه التشخيص البصري التقليدي تحدياتٍ في التمييز بين الفئات السبع الرئيسية للأمراض الجلدية نتيجة التشابه الظاهري وعدم توازن البيانات. تهدف هذه الدراسة إلى تطوير نموذج تشخيصي آلي يعتمد على شبكة عصبية تلافيفية Convolutional Neural Network (CNN) لتعزيز دقة التصنيف وتقليل الاعتماد على التقييم البشري؛ استخدمت الدراسة مجموعة بيانات HAM10000 (Human Against Machine with 10000 Training Images)، التي تضم 10,015 صورة موزعة على سبع فئات، مع معالجة عدم توازن العينات بتقنية التجاوز العشوائي للعينات Random Oversampling، ثم قُسّمت البيانات بنسبة 80% تدريب و20% اختبار، مع تخصيص 20% من بيانات التدريب للتحقق .صُمّم النموذج بتتابع من طبقات الالتفاف Conv2D والتجميع MaxPooling لاستخلاص السمات، ثم تسطيح المخرجات Flatten وربطها بطبقات كثيفة Dense للتصنيف النهائي، ودُرّب باستخدام محسّن Adam لضبط الأوزان بشكل تكيّفي ودالة الخسارة Categorical Crossentropy الملائمة للتصنيف متعدد الفئات؛ وقد حقق النموذج دقة إجمالية Accuracy بلغت 95.45%. كما أظهر تحليل الأداء باستخدام مؤشرات تقييم دقيقة مثل Precision، وRecall، و F1-Score تفوقًا ملحوظاً لمختلف الفئات، حيث تراوحت قيمة Precision بين 84% و99%، وسُجّل متوسط مرجّح لكل من Precision و F1-Score بلغ 96%. تؤكد هذه النتائج الدور الطبي الهام للنموذج المقترح كأداة ذكية داعمة للأطباء، إذ يعزز التشخيص المبكر، ويقلل الأخطاء البشرية، ويسهم في تحسين جودة الرعاية الصحية.

*Corresponding author
https://doi.org/10.63318/waujpasv3i2_29

المقدمة

الجلد هو أكبر عضو في الجسم، إذ يتفاعل مع البيئة الخارجية ويؤدي العديد من الوظائف؛ فهو يعمل كنظام دفاعي أساسي، ويحمي الأعضاء الداخلية من الإصابات الجسدية والكائنات الدقيقة الضارة، مثل البكتيريا والفيروسات، إضافةً إلى آثار الأشعة فوق البنفسجية. ويتكوّن الجلد من ثلاث طبقات مميزة: البشرة، والأدمة وتحت الجلد، حيث تسهم كلّ منهما في أداء وظائفه المعقدة. ومع ذلك، فإن الاضطرابات في العمليات الخلوية مثل الطفرات التي تسبب تدميرًا غير طبيعيًا للحمض النووي، قد تؤدي إلى تكاثر الخلايا غير المنضبط وتطور سرطان الجلد [1].

يُعتبر سرطان الجلد أحد أكثر أنواع السرطان شيوعًا وخطورة في العقد الأخير [2،3]. وتنقسم أنواعه إلى خبيثة مثل: التقرن الشعاعي (AKIEC)، وسرطان الخلايا القاعدية (BCC) والورم الميلانيني (MEL)، وأخرى حميدة مثل: التقرن الحميد (BKL)، الورم الليفي الجلدي (DF)، الآفة الوعائية (VASC) والوحمة الصبغية (NV) [3]. ويُعتبر التشخيص المبكر عنصرًا أساسيًا لرفع دقة التشخيص وتقليل معدل الوفيات [4]، تُعَدّ الخزعة وتنظير الجلد من أكثر الطرق التقليدية استخدامًا لتشخيص سرطان الجلد [3]. ورغم شيوع الخزعة بين الأطباء، إلّا أ نها إجراء جراحي مولم ويستغرق وقتًا طويلًا [2]. أما تنظير الجلد، فهو أسلوب غير جراحي يعتمد على عدسة مكبرة لفحص التغيرات الجلدية، لكنه يتطلب خبرة عالية ومراقبة دقيقة مما يجعله بطيئًا وغير عملي مع تزايد أعداد المرضى [3]. ومن أبرز عيوب الطرق التقليدية أنها تقتصر غالبًا على تصنيف الأمراض الجلدية إلى فئتين رئيستين: حميدة وخبيثة، دون القدرة على التمييز الدقيق بين الفئات الفرعية المختلفة استنادًا فقط إلى الخصائص المرئية [5]. كما هو موضح في الشكل (1).



الشكل 1: صور للأمراض الجلدية في حالات مختلفة [5]

لذلك، يُعَدّ بناء نظام تشخيص بمساعدة الحاسوب Computer Assisted Diagnosis (CAD) غير جراحي لتصنيف أمراض الجلد أمر ضروريًا وفعّالًا، إذ يساهم في تقدم المجال الطبي ويقلل معدل الخطأ في التشخيص. كما يلعب الذكاء الاصطناعي Artificial Intelligence (AI) دورًا مهمًا في الكشف عن سرطان الجلد، ولا سيما من خلال تقنيات التعلم العميق Deep Learning (DL) والتعلم الآلي Machine Learning (ML) [6].

يُعتبر التعلم العميق (DL) أحد أقوى تقنيات التعلم الآلي وأكثرها استخدامًا للتعرف على الصور وتصنيفاتها [7]، إذ يُمكّن من تحسين القدرات التشخيصية لمقدمي الرعاية الصحية، مع توفير إمكانية الكشف المبكر وبدء العلاج في الوقت المناسب [8]. جاء هذا البحث بدافع الاستفادة من إمكانيات الذكاء الاصطناعي، وخاصة التعلم العميق، في تطوير أداة طبية تساهم في تحسين جودة الرعاية الصحية، وتقديم حل تقني لمشكلة تصنيف سرطان الجلد بدقة، مع مراعاة التحديات الواقعية مثل تنوع الفئات وقلة العينات لبعضها. ويهدف البحث إلى بناء نموذج فعال يعتمد على تقنيات التعلم العميق لتصنيف سرطان الجلد إلى سبع فئات مختلفة باستخدام الصور الجلدية من مجموعة بيانات HAM10000 [9] ، مما يسهم في دعم القرار الطبي ويحسن دقة التشخيص، مع التركيز على معالجة تحديات عدم توازن الفئات وتحقيق أداء عالٍ في تصنيف كافة الأنواع دون تحيز.

خضعت الصور لعمليات معالجة مسبقة لتوحيد أبعادها ومعالجة مشكلة عدم توازن الفئات باستخدام أساليب محسنة تحافظ على تنوع العينات. تم بناء النموذج باستخدام الشبكة العصبية التلافيفية Convolutional Neural Network (CNN)، وهي خوارزمية تعلم عميقة تتكون من عدة خوارزميات عصبية تكتشف تلقائيًا خصائص البيانات المدخلة قبل تطبيق المعرفة لتطوير التدخلات [10]، وقد تم اختيارها لقدرتها على التعامل مع مجموعات البيانات الضخمة وتقديم دقة عالية في تطبيقات معالجة الصور والاشارات [7،11]. بعد ذلك، تم تدريب النموذج واختباره باستخدام مجموعة منفصلة من البيانات وتحليل أدائه بناءً على مقاييس دقيقة تشمل الدقة الكلية Accuracy ، الضبط Precision ، الاستدعاء Recall و F1-Score لكل فئة.

مقارنة بين مجموعات بيانات سرطان الجلد

يوضح الجدول (1) خصائص بيانات سرطان الجلد المختلفة. يتطلب إنشاء نظام تصنيف موثوق وقوي لسرطان الجلد وجود مجموعة متنوعة من البيانات مع جميع أنواع الصور الجلدية. ومع ذلك تتزايد الحاجة إلى موارد التصوير الطبي في الأوساط الأكاديمية وأصبحت الكثير من مجموعات بيانات أمراض الجلد متاحة للجمهور وهذه البيانات هي [12]:

1. **مجموعة بيانات PH^2:** تم انشاء هذه المجموعة بواسطة [13]، لدعم البحث في أساليب التصنيف والتجزئة. تحتوي على 200 صورة تنظير جلد ملونة (768*560) لثلاثة أنواع من أمراض الجلد، بما في ذلك الشامات الشائعة والشامات غير النمطية والأورام الميلانينية. بالإضافة إلى ذلك، تحتوي على شروح طبية شاملة، مثل نتائج تجزئة الآفات والتشخيص المرضي.
2. **مجموعة بيانات MED-NODE:** تم جمعها بواسطة قسم الأمراض الجلدية في المركز الطبي الجامعي في جرونينجن University Medical Center Groningen (UMCG) والتي تحتوي على 170 صورة رقمية لمرض الميلانوما وتستخدم للكشف عن سرطان الجلد باستخدام الصور المجهرية.
3. **مجموعة بيانات HAM10000:** جمعت بواسطة التعاون الدولي لتصوير الجلد International Skin Imaging Collaboration (ISIC) لحل مشكلة اختلال توازن البيانات ومحدوديتها في مجموعات بيانات أمراض الجلد. وتحتوي هذه المجموعة على 10,015 صورة

الجدول1: خصائص مجموعات بيانات أمراض الجلد المختلفة [12]

مجموعة البيانات	عدد الصور	نمط الصورة	عدد أنواع الأمراض	تنسيق الصورة	سنة النشر	هدف النشر
PH²	200	Dermoscopic	3	bmp	2013	تسهيل تطوير أنظمة التشخيص بمساعدة الكمبيوتر في تقسيم وتصنيف سرطان الجلد
MED-NODE	170	Macroscopic	2	jpg	2015	بناء وتقسيم نظام MED-NODE للكشف عن سرطان الجلد باستخدام الصور الجلدية
HAM10000	10,015	Dermoscopic	7	jpg	2018	لمعالجة الحجم الصغير والتنوع غير الكافي للصور في مجموعات بيانات سرطان الجلد
Derm7pt	2000	Dermoscopic Structured data	15	jpg	2018	كقاعدة بيانات لتحليل قائمة مكونة من سبع نقاط للأورام الخبيثة في الجلد
BCN20000	19,424	Dermoscopic	9	jpg	2019	تستخدم لتحليل آفات سرطان الجلد في الأماكن التي يصعب تشخيصها مثل الأظافر والأغشية المخاطية
ISIC Archive	>13,000	Dermoscopic	9	jpg	-2016 2020	لتقليل معدل الوفيات الناجمة عن سرطان الجلد مع تعزيز تطوير واستخدام التصوير الرقمي للجلد

جلدية لسبعة أمراض تمثيلية لآفات الجلد وسيتم تقديم وصف لهذه في الفئات في القسم التالي.

4. **مجموعة بيانات Derm7pt:** تحتوي على ما يقرب من 2,000 صورة ملونة سريرية وجلدية لأمراض الجلد بالإضافة إلى معلومات منظمة لتدريب وتقييم أنظمة التصميم بمساعدة الكمبيوتر وهي بمثابة قاعدة بيانات لتحليل نتائج التنبؤ بقائمة التحقق من الأورام الخبيثة المكونة من سبعة نقاط للآفات الجلدية.
5. **مجموعة بيانات BCN 20000:** تتكون هذه المجموعة من 5,583 صورة آفة جلدية و19,424صورة جلدية منظارية مأخوذة باستخدام منظار جلدي عالي الدقة تم جمعها جميعًا في عامي 2010 و2016. في الوقت نفسه، استخدم جامع البيانات مجموعة متنوعة من تقنيات الرؤية الحاسوبية لإزالة الضوضاء والخلفية والتدخلات الأخرى من الصور. واخيرًا تمت مراجعتها بعناية من قبل الخبراء لضمان دقة التشخيص.
6. **مجموعة بيانات ISIC:** لتقليل وفيات سرطان الجلد مع تعزيز تطوير واستخدام التصوير الرقمي للجلد، أنشأ التعاون الدولي لتصوير الجلد (ISIC) مجموعة بيانات لأمراض الجلد متاحة للعامة. يضم أرشيف هذه المجموعة حاليًا أكثر من 13,000 صورة تمثيلية للجلد وقد تم فحصها من قبل خبراء لضمان جودة الصور.

الدراسات السابقة

تعرض هذه الفقرة أبرز الدراسات السابقة التي تناولت تطبيق خوارزميات التعلم العميق في تصنيف صور سرطان الجلد، وذلك بالاعتماد على مجموعات بيانات متنوعة. وتركز هذه الدراسات على النماذج المستخدمة، طرق المعالجة المسبقة، وأداء النماذج في تصنيف الآفات الجلدية.

1. اكتشاف سرطان الجلد اعتمادًا على صورالآفات الجلدية باستخدام التعلم العميق

قام Gouda وآخرون (2022) بإجراء دراسة تهدف إلى تطوير نظام ذكي يعتمد على تقنيات التعلم العميق للكشف المبكر عن سرطان الجلد اعتمادًا على صور الآفات الجلدية. اعتمد الباحثون في دراستهم على قاعدة بيانات ISIC 2018، التي تحتوي على مجموعة ضخمة من صور الأورام الجلدية المصنفة إلى حميدة وخبيثة. وقد قاموا بتطبيق مجموعة من تقنيات المعالجة المسبقة للصور مثل تحسين جودة الصور باستخدام تقنية ESRGAN، وتكبير البيانات من خلال عمليات التدوير، والانعكاس، وتغيير الحجم، بالإضافة إلى تطبيع قيم الصور لجعلها ضمن نطاق موحد. اعتمدت الدراسة على عدة نماذج من الشبكات العصبية التلافيفية (CNN)، منها نموذج تم تصميمه خصيصًا للدراسة، بالإضافة إلى نماذج مدربة مسبقًا مثل ResNet50، وInceptionV3، وInception-ResNet. تم استخدام مجموعة من 3533 صورة، حيث تضمنت 1760 صورة لآفات حميدة و1773 صورة لآفات خبيثة، وقد تم تقسيم البيانات إلى مجموعات تدريب واختبار بنسبة 80% و20% على التوالي، أظهرت نتائج الدراسة أن نموذج InceptionV3 حقق أفضل أداء بدقة بلغت 85.7%، في حين أن النموذج المصمم من طرف الباحثين بلغ دقته 83.1%. كما تم تقييم أداء النماذج باستخدام عدة مقاييس منها الدقة، والاسترجاع، F1-Score ، ومنحنى ROC-AUC. وقد أظهرت النتائج أن النموذج المقترح يوفر أداءً فعالًا، حيث أظهر قدرة على التمييز بين الأورام الحميدة والخبيثة بدرجة تقارب أداء أطباء الجلد ذوي الخبرة [14].

2. التصنيف متعدد الفئات لسرطان الجلد باستخدام الشبكة العصبية التلافيفية CNN

قدم Shobha وآخرون (2022) نموذج لتصنيف سرطان الجلد متعدد الفئات باستخدام الشبكة العصبية التلافيفية (CNN) مدعومًا بتقنية تعزيز البيانات Data Augmentation على مجموعة بيانات HAM10000. واشتملت المعالجة المسبقة استبدال القيم المفقودة وتصغير ابعاد الصور لتقليل الحسابات، وتم تقسيم البيانات إلى تدريب واختبار وتحقق حيث تحتوي بيانات التدريب على 6696 صورة والاختبار 2481 صورة والتحقق 744 صورة. إضافة إلى تقنيات تعزيز مثل التدوير والتكبير والانعكاس الأفقي والرأسي لموازنة الفئات، وتم تقييم النموذج باستخدام مصفوفة التشتت Confusion Matrix والدقة Accuracy والضبط Precision والاستدعاء Recall و F1-Score وحقق النموذج دقة اختبار بلغت 77.91% ودقة تدريب بلغت 86.54% مع متوسط دقة مرجحة 0.77 ودقة ماكرو 0.62 مما يدل على كفاءته في تصنيف سبعة أنواع من آفات الجلد [15].

3. تحسين اكتشاف سرطان الجلد باستخدام تحسين دقة الصور والشبكات العصبية الالتفافية

قام Lembhe وآخرون (2023) بتطوير طريقة آلية للكشف عن سرطان الجلد باستخدام معالجة الصور والتعلم الآلي. استخدموا تقنيات Image Super-Resolution (ISR) لزيادة دقة صور آفات الجلد من مجموعة بيانات ISIC التي تم تقسيمها بنسبة 20% للتدريب و80% للاختبار، والتي تحتوي على صور حميدة وخبيثة. طبقّوا شبكات عصبية تلافيفية (CNN) مثل VGG16 وResNet وInceptionV3 لتصنيف الصور إلى فئتين (تصنيف ثنائي) أي حميد وخبيث. هدفت الدراسة إلى تحسين دقة الكشف المبكر عن سرطان الجلد، وهي مشكلة صعبة بسبب تشابه الصور الطبية وحاجة تخزين الصور عالية الدقة. أظهرت النتائج أن ISR حسّنت دقة CNN، وحقق InceptionV3 أعلى دقة 91.26٪ [16].

4. تصنيف سرطان الجلد باستخدام NASNet

قام Rahman وآخرون (2024) بدراسة تهدف إلى تطوير نموذج فعال لتصنيف سرطان الجلد باستخدام تقنيات التعلم العميق، وذلك من خلال تطبيق شبكة NASNet المحسّنة على صور الآفات الجلدية. وقد سعى الباحثون إلى تصنيف الآفات إلى فئتين رئيسيتين: الميلانوما (الخبيثة) وغير الميلانوما (الحميدة)، مع التركيز على تحسين أداء النموذج عند التعامل مع بيانات طبية غير مكتملة أو غير متوازنة. استخدموا في دراستهم قاعدة بيانات تتألف من 2637 صورة لآفات جلدية مأخوذة من مجموعة بيانات ISIC، حيث تضمنت 1197 صورة لأورام خبيثة و1440 صورة لآفات حميدة، وتم تقسيمها إلى مجموعات تدريب 80% واختبار 20%. اعتمدت الدراسة على نموذج NASNet Mobile وNASNet Large، وتم تعزيزها من خلال تقنيات معالجة مسبقة شاملة مثل تغيير الحجم، التدوير، الإزاحة الأفقية والعمودية، والتكبير العشوائي، بالإضافة إلى التطبيع اللوني، وذلك بهدف تحسين تعميم النموذج وتقليل الإفراط في التعلّم. تم تدريب النموذج باستخدام خوارزميات تحسين مثل Adam وNadam، مع استخدام دالة الخسارة من نوع categorical cross-entropy، وتفعيل الدوال غير الخطية مثل ReLU وSigmoid. أظهرت نتائج الدراسة أن نموذج NASNet Mobile حقق دقة تصنيف بلغت 85.62%، بينما بلغت دقة NASNet Large نحو 83.98%. وقد بيّنت مؤشرات الأداء الأخرى مثل الحساسية، الدقة، ومعامل F1 أن النموذج المقترح يتفوق على العديد من النماذج الأخرى مثل ResNet، VGG، وMobileNet. كما أظهرت مقارنة شاملة أن NASNet قدّم أداءً عاليًا في تصنيف صور الجلد، مما يبرز إمكانيته كأداة دعم فعالة في التشخيص الطبي المعتمد على الصور [17].

5. كشف سرطان الجلد باستخدام التعلم العميق

قام Pathania وBehki (2024) بتقديم نظام مدعم بالتعلم العميق، وتحديدًا نموذج DenseNet 169 المدرب على مجموعة بيانات (Skin Cancer: Malignant Vs Benign) وتتكون مجموعة البيانات هذه من مجلدين يحتوي كل منهما على 1800 صورة بتنسيق JPEG لآفات الجلد بحجم 224*224 وتصنف هذه الصور إلى نوعين من الشامات: الحميدة والخبيثة. وتم معالجة هذه البيانات مسبقًا عن طريق إزالة التكرارات ومعالجة القيم المفقودة وتوحيد احجام الصور وتطبيع قيم البكسل لضمان الاتساق، تم تقسيم هذه البيانات إلى 80% تدريب و20% اختبار وتم زيادة البيانات لإثراء التنوع في مجموعة التدريب وتعزيز قدرة النموذج على تعميم قوته واشتملت هذه التقنيات: الدوران، الإزاحة، القص والتكبير. وتم تقييم النموذج من خلال عدة مقاييس بما في ذلك الدقة Accuracy ، الضبط Precision ، الاستدعاء Recall، F1-Score ، AUC-ROC ومصفوفة التشتت Confusion Matrix وبلغت الدقة التي حققها نموذج DeseNet169 في هذه الدراسة 89.7% في تصنيف آفات الجلد باستخدام صور التنظير الجلدي وبذلك تهدف هذه الدراسة إلى توفير أداة مساعدة في المختصين في الرعاية الصحية لتشخيص سرطان الجلد بشكل أسرع وأكثر دقة [8].

6. تحسين تشخيص سرطان الجلد باستخدام الشبكات العصبية التلافيفية العميقة: نهج تحسين متعدد النماذج

تهدف هذه الدراسة التي أجريت من قبل Nataragan وKumar (2024) إلى تحسين دقة تشخيص سرطان الجلد باستخدام شبكات عصبية التفافية عميقة (CNN) وتقنية النماذج المجمعة Ensemble المعتمدة على التعلم بالنقل Transfer Learning. اعتمدت الدراسة على مجموعة بيانات ISIC2019 والتي تحتوي على 25,335 صورة تغطي سبعة أنواع من الأورام الجلدية حيث تم تقسيم البيانات إلى 80% للتدريب و10% للتحقق و 10% للاختبار، مع تطبيق تقنيات تعزيز الصور Data Augmentation لزياد فعالية النموذج. شملت الدراسة ايضًا عمليات معالجة مسبقة للصور مثل القص والتقليم لتحسين جودة الادخال. واستخدمت الدراسة نماذج (CNN) مسبقة التدريب وتشمل : InceptionV3 ، ReseNet 50 ،VGG19 ،VGG16 و Xception وتم دمج بعضها لتحسين الأداء. وأظهرت النتائج أن نموذج VGG16مع Xception حقق دقة اختبار بلغت 82.37% بينما حقق VGG19مع inceptionN3 أقل قيمة لفقدان الاختبار (Loss). خلصت الدراسة إلى أن دمج النماذج يمكن أن يساهم في تحسن دقة التصنيف لكن اشارت إلى أن عوامل مثل تنوع البيانات وضبط المعلمات تؤثر على النتائج [4].

من خلال مقارنة مجموعات البيانات المختلفة في الجدول (1) واستعراض الدراسات السابقة التي تناولت استخدام تقنيات التعلم العميق في تشخيص سرطان الجلد، تبيّن أن أبرز التحديات تكمن في محدودية حجم البيانات ونقص تنوعها. حيث أن معظم الصور في مجموعات بيانات أمراض الجلد هي صور جلدية، مما لاتزال الصور السريرية والصور النسيجية نادرة. علاوة على ذلك تحتوي معظم بيانات أمراض الجلد على عدد قليل نسبيًا من الصور مقارنةً بمجموعات بيانات الصور الصبغية مما يشكل تحديات معينة لمهام تصنيف سرطان الجلد [12].

كذلك من أبرز التحديات التي تواجه تدريب الشبكات العصبية للتشخيص الآلي لأمراض الجلد الصبغية صغر حجم مجموعات البيانات ونقص تنوعها، وقد تم حل هذه المشكلة من خلال اصدار مجموعة بيانات HAM10000 [18]. تعد مجموعة بيانات HAM10000 [9] الأفضل مقارنةً بمجموعات البيانات الأخرى مثل PH^2 و ISIC لأنها تضم عدد كبير من الصور (10,015) صورة من مصادر متعددة وتغطي جميع الأمراض الجلدية على عكس PH^2 التي تحتوي على عدد محدود من الصور و ISIC التي تركز بشكل مفرط على سرطان الجلد Melanoma والشامات، فإن HAM10000 تقدم تمثيلًا أكثر واقعية وشمولًا مما يجعلها مرجعًا قياسيًا في أبحاث تصنيف الصور الجلدية [18]. لذلك تم استخدام مجموعة بيانات

HAM10000 في هذا البحث لما توفره من تنوع كبير وجودة عالية مما يجعلها مناسبة لتدريب وتقييم نماذج تصنيف الآفات الجلدية بدقة وموثوقية.

نموذج الشبكة العصبية التلافيفية Convolutional Neural Networks (CNN)

تشبه الشبكات العصبية التلافيفية Convolutional Neural Networks (CNN) الشبكة العصبية الصناعية التقليدية Artificial Neural Networks (ANN) من حيث أنها تتكون من خلايا عصبية تعمل على التحسين الذاتي من خلال التعلم. يتمثل الاختلاف الوحيد الملحوظ بين الشبكة العصبية التلافيفية والشبكة العصبية الصناعية التقليدية في أن شبكات (CNN) تستخدم بشكل أساسي في مجال التعرف على الأنماط داخل الصور. وهذا يسمح لنا بترميز ميزات خاصة بالصور في البنية مما يجعل الشبكة العصبية أكثر ملائمة للمهام التي تركز على الصور مع تقليل المعلمات المطلوبة لإعداد النموذج بشكل أكبر [19]. كما أنها تعد واحدة من أفضل خوارزميات التعلم العميق حيث أنها تظهر اداءً مثاليًا في المهام المتعلقة بتجزئة الصور واكتشافها وتصنيفها [20].

تتكون البنية الرئيسية للشبكة العصبية التلافيفية (CNN) من طبقة الالتفاف Convolutional Layer وطبقة التجميع Pooling Layar وطبقة التنشيط غير الخطية Nonlinear Activation Layer وطبقة الاتصال الكامل Fully Connected Layer والشكل (2) يوضح البنية العامة للشبكة العصبية التلافيفية (CNN) . عمومًا تعالج الصور مسبقًا، ثم تدخل إلى الشبكة عبر طبقة الادخال Input Layer وتعالج بواسطة عدة طبقات وهي طبقات الالتفاف وطبقات التجميع مرتبة بالتناوب، تم تصنف حسب طبقة الاتصال بالكامل [21].

أنواع الطبقات في الشبكات العصبية التلافيفية:

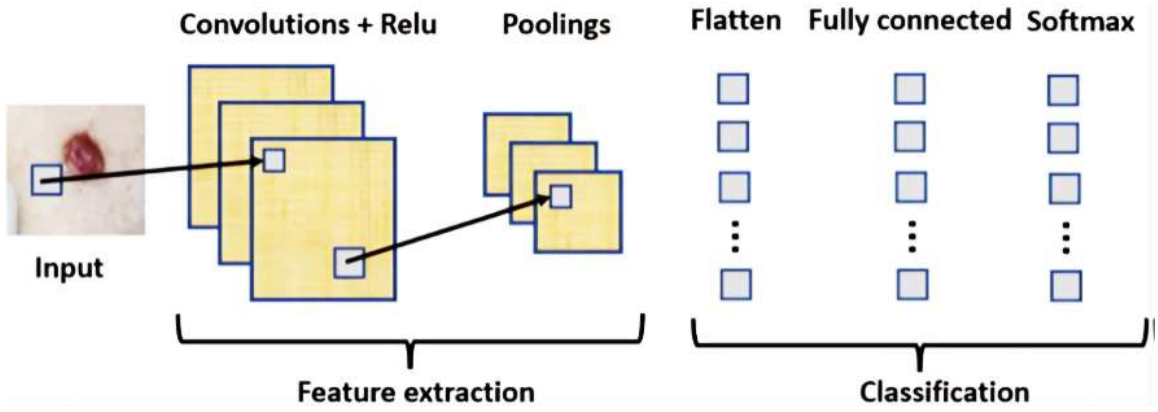
تتكون الشبكات العصبية التلافيفية (CNN) من مجموعة من الطبقات الأساسية ولكل منها دور محدد في عملية معالجة الصور وتصنيفها. وفيما يلي نظرة عن هذه الطبقات [21]:

- **طبقة الالتفاف Convolutional Layer:** تستخدم لاستخلاص الميزات الأساسية من الصور عبر تطبيق مرشحات على مناطق محلية من البيانات، وتتميز بمشاركة الأوزان والاتصال المحلي، مما يقلل من عدد المعاملات ويزيد من الكفاءة.
- **طبقة التجميع Pooling Layar:** تعمل على تقليل الأبعاد المكانية للبيانات وتقليل الحمل الحسابي، وتساعد في تعزيز ثبات النموذج ضد التحولات الطفيفة. أكثر أنواعها شيوعًا التجميع الأقصى والمتوسط.
- **طبقة التفعيل Activation Layer:** تضيف لا خطية على الشبكة، مما يمكنها من تعلم علاقات أكثر تعقيدًا تعد ReLU أكثر الدوال استخدامًا بسبب بساطتها وفعاليتها في تسريع التدريب وتقليل مشاكل التدرج.
- **الطبقة المتصلة بالكامل Fully Connected Layer:** تربط جميع الخلايا العصبية من الطبقات النهائية وتستخدم في تصنيف الميزات المستخلصة وغالبًا ما تستخدم معها طبقة SoftMax للحصول على احتمالات الفئات.

بالإضافة إلى أنواع الطبقات لبنية CNN يتم تحقيق التصنيف النهائي من طبقة الإخراج Output Layer التي عادة ما تكون الطبقة الأخيرة من طبقة Fully Connected Layer ، تؤثر دالة الخسارة Loss Function على أداء بنية (CNN) حيث انها تقوم بقياس الفرق بين المخرجات المتوقعة والمخرجات الحقيقية وتوجه عمليات التعلم. في تصنيف الصور عادة تستخدم دالة Cross Entropy مع SoftMax لتحقيق دقة عالية [21]. ولتقليل دالة الخسارة وتحسين الأوزان تستخدم خوارزمية التحسين Optimizer اثناء التدريب، ومن أبرزها انحدار التدرج العشوائي Stochastic Gradient Descent (SGD) وتقدير العزم التكيفي Adaptive Moment Estimation (Adam) حيث تسهم في تسريع التقارب وتحسين الأداء العام للنموذج [21].

المواد وطرق العمل

اعتمدت هذه الدراسة على التعلم العميق لتصنيف صور سرطان الجلد إلى سبع فئات تشخيصية باستخدام الشبكات العصبية الالتفافية Convolutional Neural Network (CNN). تتضمن المراحل الأساسية: جمع البيانات، المعالجة المسبقة، تصميم بنية الشبكة، التدريب، الاختبار والتقييم، يوضح الشكل (3) تسلسل هذه المراحل.



الشكل 2 : البنية العامة للشبكة العصبية التلافيفية (CNN) [22]

الشكل 3: المراحل الأساسية لخطوات العمل المقترحة.

في البداية يجب الإشارة إلى أنه تم الاستناد على الكود الموجود في [23] وتطويره من خلال إضافة بعض مقاييس الأداء الأخرى مثل مصفوفة التشتت Confusion Matrix وحساب الضبط Precision والاستدعاء Recall وF1-score لكل فئة وتم عرض هذه المقاييس في تقارير التصنيف Classification Reports ومنحنيات تغير الدقة والخسارة.

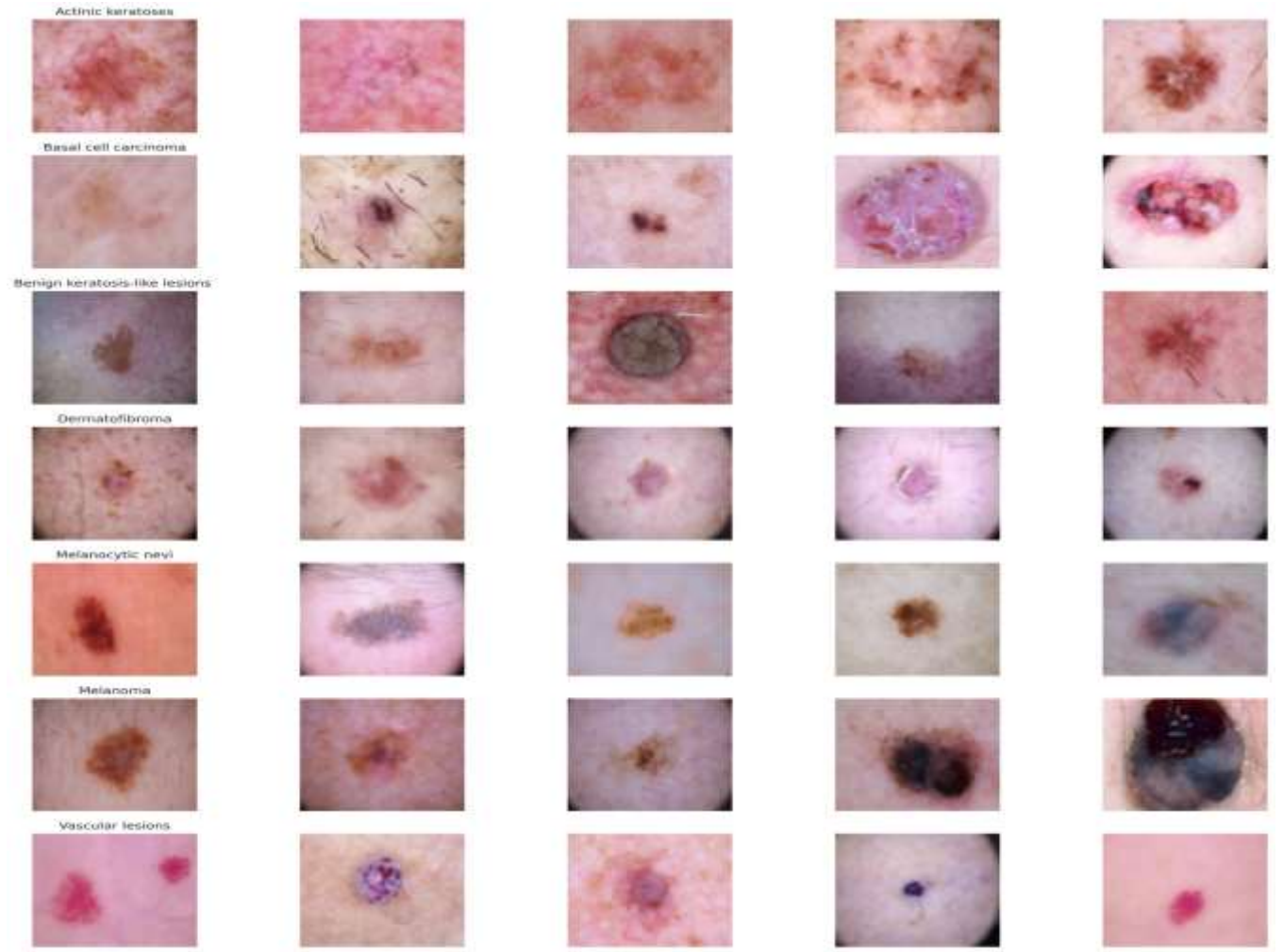
تم تنفيذ الكود البرمجي على جهاز كمبيوتر شخصي بالمواصفات التالية: نوع الحاسوب hp نظام التشغيل Windows10مزود بمعالج™ Intel ® core i7-8550U CPU @ 1.80 GHZ 1.99GHZ من الجيل الثامن، وذاكرة وصول عشوائي (RAM) بسعة 16.0 غيغابايت، والإصدار 64 بت وتم استخدام دفتر جوبيتر Jupyter Notebook بإصدار 7.2.2 ضمن منصة Anaconda بإصدار 2.6.3 لتشغيل النموذج. وكان إصدار المكتبات المستخدمة: pandas 2.2.3, matplotlib 3.10.0 Seaborn 0.13.2, numpy 1.26.4, imblearn 0.12.3, sklearn 1.6.1, tensorflow 2.10.0, keras 2.10.0 وفيما يلي شرح لكل مرحلة من مراحل المنهجية المتبعة بالتفصيل:

1. مجموعة بيانات HAM10000

تضم هذه المجموعة صورًا مختلفة لسرطان الجلد، ويتضمن استخدام اساليب التعلم العميق كمية كبيرة من البيانات لإعطاء نتائج مرضية. حيث تتكون مجموعة البيانات هذه من 10,015 صورة تنظير جلدي [6]. وتتكون من صور تصور أمراض الجلد وتصنف هذه الأمراض إلى سبع فئات مختلفة بناءً على خصائصها [6]، والشكل (4) يعرض عينات كل فئة.

يشمل تصنيف الأورام والأمراض الجلدية سبعة أنواع، وفيما يلي تعريف لكل نوع [4]:

- **سرطان الخلايا الظاهرية Actinic Keratosis Intraepithelial (Carcinoma-AKIEC): هو** نوع شائع من سرطان الخلايا الحرشفية ويعد الكشف المبكر عنه أمر بالغ الأهمية للحفاظ على الصحة.



الشكل 4: خمس عينات من كل فئة من مجموعة بيانات HAM10000 [6].

- **سرطان الخلايا القاعدية (Basal Cell Carcinoma-BCC):** يعد أكثر أشكال سرطان الجلد الخبيث شيوعًا ويُظهر معدل نمو بطيء ولديه ميل منخفض للانتشار. ومع ذلك، يمكن أن يؤدي إلى تلف موضعي كبير في الأنسجة إذا ترك من دون علاج.
- **التقرن الحميد (Benign Keratosis-BKL):** تشمل هذه الفئة التقرن الدهني وهو نمو حميد قد يشبه الثآليل.
- **الورم الليفي الجلدي (Dermatofibroma-DF):** وهو عقيدات ليفية غير سرطانية تظهر عادة على أدمة الأطراف السفلية. تمتلك هذه الأجسام المعنية عمومًا بنية صلبة ولديها القدرة على اظهار التلون.
- **الورم الميلانيني (Melanoma-MEL):** هو شكل خبيث للغاية من سرطان الجلد الذي يمكن أن ينشأ من شامات موجودة مسبقًا أو يظهر تلقائيًا على الجلد غير المصاب. يزداد احتمال حدوث النقائل إذا لم يتم اكتشاف الحالة في مراحلها المبكرة.
- **الشامات الميلانينية (Melanocytic Nevus-NV):** وهي آفات جلدية مصطبغة تعرف باسم الشامات. وهي عادة ما تكون غير سرطانية بطبيعتها ومع ذلك فإن بعض المتغيرات لديها القدرة على التطور إلى سرطان الجلد.
- الآفات الوعائية (**Vascular Lesion-VASC**): تشمل هذه الفئة مجموعة من الأورام غير السرطانية والسرطانية التي تنشأ من الأوعية الدموية بما في ذلك الأورام الوعائية وساركوما كابوزي Kaposi 's Sarcoma.

تحتوي مجموعة بيانات HAM10000 على صور عالية الدقة، مرفقة ببيانات وصفية وتشخيصية. مجموعة البيانات المستخدمة في هذا البحث تم تنزيلها من [9] واستخدامنا ملفين رئيسيين هما:

- HAM10000_metadata.csv ويحتوي على معلومات المرضى والآفات.
- hmnist_28_28_RGB.csv والذي يتضمن بيانات الصور في صيغة جدولية (28×28 بكسل، بصيغة RGB)إلى جانب التصنيفات.

تتسم مجموعة بيانات HAM10000 بمشكلة عدم التوازن حيث تحتوي بعض الفئات على حالات أقل بكثير مقارنةً بفئات أخرى، ويوضح الشكل (5) هذه الاختلال بشكل بصري من خلال عرض التوزيع غير المتساوي للبيانات عبر مختلف الفئات [6]. وللتغلب على هذه المشكلة تم استخدام تقنية التجاوز العشوائي لعينات Random Oversampling، والتي سيتم تناولها لاحقًا ضمن مرحلة المعالجة المسبقة للبيانات.

2. المعالجة المسبقة للبيانات

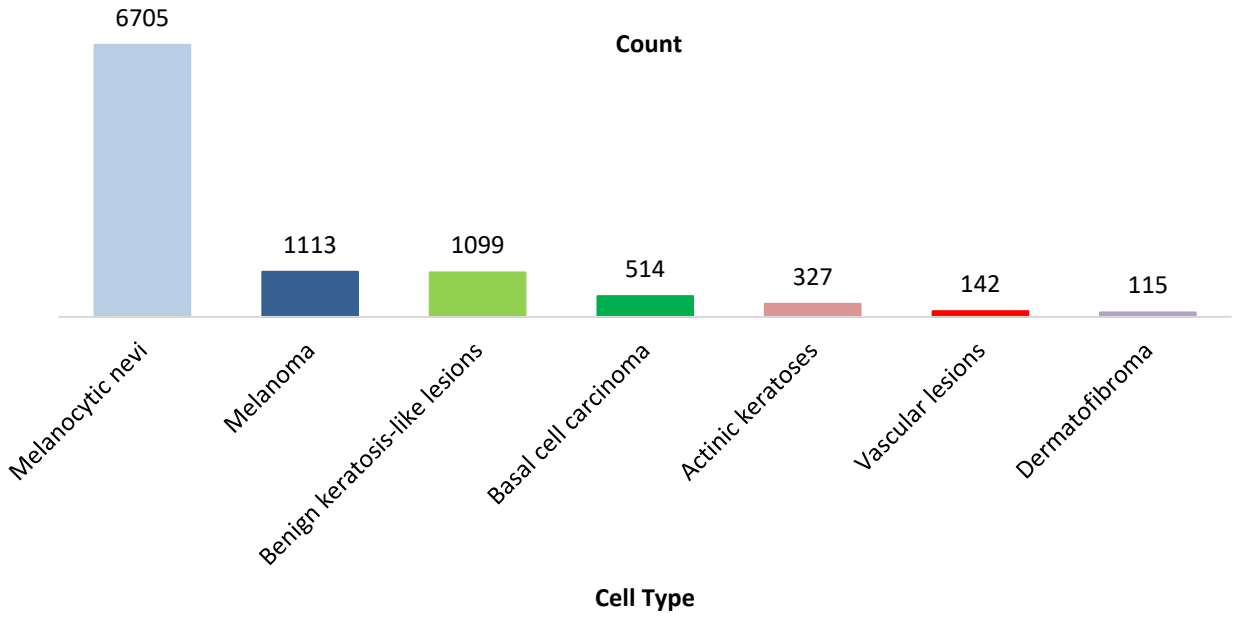
تم تطبيق مجموعة من الخطوات المهمة لتحضير البيانات الخام لتكون مناسبة لتدريب النموذج العميق، وهي كالتالي:

- **إعادة تشكيل الصور**

تمثل الصور في الملف hmnist_28_28_RGB.csv صيغة مسطّحة flattened مكوّنة من 2352 قيمة عددية لكل صورة، نتيجة ضرب الأبعاد 3×28×28لإدخالها في نموذج CNN أُعيدت تشكيل كل صورة إلى المصفوفة ثلاثية الأبعاد بحجم 3×28×28 فتتحوّل مجموعة البيانات إلى مصفوفة رباعية الأبعاد بالشكل (عدد_الصور، 28، 28، 3)، مما يسمح للنموذج بالتعامل مع العرض والارتفاع والقنوات اللونية واستخراج الأنماط البصرية بكفاءة.

- **التوحيد القياسي (Standardization)**

هو أسلوب قياس يستخدم في خطوة تحويل البيانات. فهو يحول مجموعة البيانات بطريقة ينتج عنها متوسط يساوي صفر وانحراف معياري يساوي واحد [24]، ويتم حسابه بالمعادلة (1):



الشكل 5: وصف كل فئة في مجموعة بيانات HAM10000 [6]

$$Z = \frac{x - \mu}{\delta} \quad (1)$$

حيث يشار إلى Z بالقيمة القياسية، و x هي القيمة الأصلية لسمة محددة، و μ، δ هما المتوسط الانحراف المعياري للبيانات على التوالي[24]. ولحماية نموذج (CNN) من تأثير التوزيع غير المتوازن للقيم اللونية، تم في هذا البحث إجراء التقييس(Standardization) باستخدام المتوسط والانحراف المعياري للبيانات هذا يساعد على تسريع عملية التدريب وتحسين استقرار النموذج.

- **معالجة عدم توازن الفئات(Class Imbalance)**

أظهرت البيانات المستخدمة في هذه الورقة وجود اختلال في توزيع الفئات، حيث كانت بعض أنواع سرطان الجلد ممثلة بعدد قليل من العينات مقارنة بغيرها. ولتفادي انحياز النموذج نحو الفئات الأكبر، تم تطبيق تقنية التجاوز العشوائي للعينات Random Oversampling ما رفع عدد العينات الإجمالي إلى 46,935. ويعرف التجاوز العشوائي بأنه إحدى الطرق المستخدمة لمعالجة مشكلة اختلال التوازن بين الفئات في مجموعات البيانات، ويُقصد به زيادة عدد العينات في الفئات الأقل تمثيلاً (الفئات الصغرى) من خلال تكرار بعض العينات الموجودة فيها بشكل عشوائي. يتم ذلك بهدف موازنة توزيع الفئات في بيانات التدريب، مما يساعد نماذج التعلم العميق، وخصوصاً الشبكات العصبية، على تعلم التمثيل الصحيح لكافة الفئات دون أن يطغى تأثير الفئات الكبرى على النموذج. يُعد التجاوز العشوائي من أبسط وأكثر الأساليب استخداماً، وقد أثبت فعاليته في تحسين دقة التصنيف، خاصة في سيناريوهات التعلم العميق، دون أن يؤدي بالضرورة إلى الإفراط في التعلّم Overfitting، كما هو الحال في بعض النماذج التقليدية [25].

3. تقسيم البيانات(Data Splitting)

بعد إتمام الخطوات السابقة، تم تقسيم البيانات إلى ثلاث مجموعات رئيسية لضمان تدريب النموذج بشكل فعّال وتقييم أدائه بدقة. في البداية، تم تقسيم البيانات بنسبة 80% للتدريب و20% للاختبار. بعد ذلك، تم تخصيص 20% من مجموعة التدريب لاستخدامها كمجموعة تحقق (Validation)أثناء عملية التدريب، وبهذا يكون التوزيع النهائي للبيانات كالتالي: 64% للتدريب، 16% للتحقق، و20% للاختبار. يهدف هذا التقسيم إلى تحقيق توازن بين تدريب النموذج وتقييم أدائه على بيانات لم يشاهدها أثناء التعلم.

4. بناء نموذج الشبكة العصبية التلافيفية Convolutional Neural Networks (CNN)

يوضح الشكل (6) البنية المعتمدة للشبكة العصبية الالتفافية (CNN) المستخدمة في هذا البحث. تتكون الشبكة من طبقة إدخال بأبعاد 28×28×3، تليها طبقتان Conv2D مع دالة تفعيلReLU ؛ الأولى تحتوي على 16 فلترًا بحجم 3×3 والثانية 32 فلترًا بحجم 3×3. بعد ذلك، توجد طبقة MaxPooling2D بحجم 2×2 لتقليل الأبعاد المكانية. ثم أُضيفت طبقتان Conv2D إضافيتان (32 و64 فلترًا) متبوعتان بطبقة MaxPooling2Dثانية لتجميع الميزات المستخرجة. بعد تسطيح البيانات باستخدام طبقةFlatten ، تمرر الشبكة النتائج عبر ثلاث طبقات Dense بأحجام 64، 32، و7 وحدات على التوالي، حيث تستخدم الطبقة الأخيرة دالة SoftMax لتوزيع احتمالات الانتماء إلى الفئات السبع المختلفة.

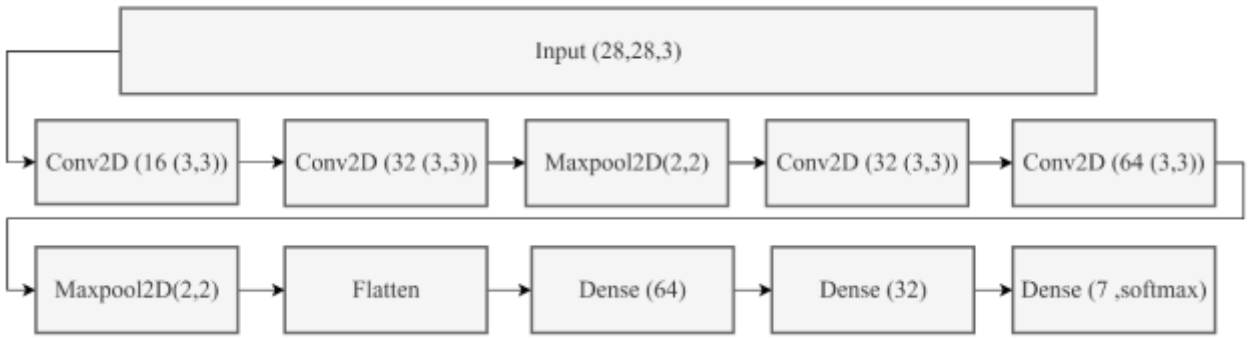
ويعرض الجدول (2) المعلمات التفصيلية لبنية الشبكة العصبية الالتفافية (CNN)المستخدمة في هذا البحث. يحتوي الجدول على أسماء الطبقات وأنواعها، أشكال مخرجات كل طبقة، وعدد المعلمات القابلة للتعلم، مما يتيح فهم حجم كل طبقة وعدد المعلمات القابلة للتعلم. يعكس هذا الجدول الهيكل التفصيلي للنموذج ويربط بين الجانب البرمجي والجانب النظري لتصميم الشبكة.

الجدول 2: معلمات النموذج CNN المستخدم في البحث

الطبقة (النوع)	شكل الإخراج	المعلمة
conv2d (Conv2D)	(None, 28, 28, 16)	448
conv2d_1 (Conv2D)	(None, 26, 26, 32)	4640
max_pooling2d (MaxPooling2D)	(None, 13, 13, 32)	0
conv2d_2 (Conv2D)	(None, 13, 13, 32)	9248
conv2d_3 (Conv2D)	(None, 11, 11, 64)	18496
max_pooling2d_1 (MaxPooling2 D)	(None, 6, 6, 64)	0
flatten (Flatten)	(None, 2304)	0
dense (Dense)	(None, 64)	147520
dense_1 (Dense)	(None, 32)	2080
dense_2 (Dense)	(None, 7)	231

5. تدريب النموذج

بلغ عدد المعاملات القابلة للتعلّم في هذا النموذج 182,663معاملًا، وتم تدريب النموذج لمدة 20حقبة epochs باستخدام المحسّن 'adam' ودالة الفقد 'Categorical_Crossentropy'، مع اعتماد الدقّة Accuracy كمقياس رئيسي للأداء. تم استخدام validation_split=0.2 لتخصيص بيانات تحقق داخلية، فيما حُفظ أفضل نموذج تلقائيًا وفقًا لقيمة val_accuracy وكان حجم الدفعة Batch Size يساوي128 ، والجدول (3) يوضح هذه المعلمات. من سجلات التدريب:



الشكل 6: بنية لشبكة العصبية الالتفافية (CNN) المستخدمة في البحث.

- بلغت أعلى قيمة للتحقق val_accuracy ≈ 97.35% عند الحقبة 19، في حين وصلت دقّة التدريب في نفس الحقبة إلى 99.53% (بفرق تقريبي 2.2%).
- سُجّلت أدنى قيمة لخسارة التحقق val_loss ≈ 0.1118 عند الحقبة 15.

الجدول 3: معلمات المحاكاة المستخدمة في البحث.

معلمات المحاكاة	القيمة
حجم الدفعة Batch Size	128
المحسن Optimizer	Adam
الحقبة Epochs	20

6. التنبؤ بصورة واحدة جديدة (اختبار النموذج)

لتقييم قدرة النموذج على التعميم، تم اختباره على صورة لم تُستخدم أثناء التدريب أو التحقق، خضعت الصورة لنفس خطوات المعالجة المسبقة بما في ذلك إعادة التحجيم وتوحيد القيم. نجح النموذج في تصنيف الصورة بدقة عالية، حيث تنبأ بفئة "vasc" بنسبة ثقة Confidence بلغت 99.9%، مما يؤكد فعاليته في التعامل مع صور جديدة لم يرها أثناء التدريب.

7. تقييم النموذج

تم تقييم أداء نموذج (CNN) المستخدم في البحث باستخدام مقاييس تقييم أهمها الدقة Accuracy. تعرف **الدقة Accuracy** بأنها نسبة الحالات المصنفة بشكل صحيح إلى اجمالي الحالات وتوفر مقياسًا عامًا لصحة النموذج [6] ، ويتم حساب الدقة من المعادلة (2):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

كذلك تم حساب الضبط Precision والاستدعاء Recall و F1-Score لكل فئة وتم عرض النتائج في تقارير التصنيف Classification Reports. ويمكن تعريف كل مقياس كما يلي:

- **الاستدعاء Recall:** يعرف ايضًا باسم الحساسية Sensitivity أو معدل الإيجابية الحقيقي True Positive Rate ويقيس قدرة النموذج على تحديد الإيجابية بشكل صحيح [6]، ويحسب من المعادلة (3):

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

- **الدقة المثالية (الضبط) Precision:** يشير إلى نسبة الحالات التي تم تحديدها بدقة على أنها إيجابية مقارنة بالحالات الإجمالية التي صنفها النموذج على أنها إيجابية [26]، ويتم تعريف الضبط بالمعادلة (4):

$$Precision = \frac{TP}{TP + Fp} \quad (4)$$

ويعمل الضبط Precision كمؤشر على موثوقية النموذج في التنبؤ بوجود حالة إيجابية، وبالتالي التأكيد على ضرورة التخفيف من النتائج الإيجابية الكاذبة، وخاصة في مجال التشخيص الطبي [26].

- **F1-Score**: وهو المتوسط التوافقي للضبط Precision والاستدعاء Recall وهو مقياس متوازن يأخذ في الاعتبار كلًا من الإيجابيات الخاطئة والسلبيات الخاطئة [6]، ويتم حسابه بالمعادلة (5) [22]:

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (5)$$

حيث أن:

TP: (True Posative) الحالة التي تكون فيها البيانات موجبة ومصنفة بشكل صحيح على أنها موجبة.

TN: (True Negative) الحالة التي تكون فيها البيانات سالبة ومصنفة بشكل صحيح على أنها سالبة.

FP: (False Posative) الحالة التي تكون فيها البيانات سالبة ومصنفة بشكل غير صحيح على أنها موجبة.

FN: (False Negative) الحالة التي تكون فيها البيانات موجبة ومصنفة بشكل غير صحيح على أنها سالبة.

ايضًا تم عرض مصفوفة الالتباس Confusion Matrix والتي توفر معلومات حول نموذج (CNN) المستخدم عبر الفئات. تعرض Confusion Matrix التنبؤات الإيجابية (True Posative) والسلبية الحقيقية (True Negative) والايجابيات الكاذبة (False Posative) والسلبيات الكاذبة (False Negative)، حيث توفر رؤية مفصلة لأنواع الأخطاء التي يرتكبها النموذج ويستخدم كركيزة أساسية لحساب مقاييس الأداء المختلفة [8].

في هذا البحث تم تقييم أداء النموذج من خلال عدة مراحل متكاملة لضمان دقة التنبؤ وفعاليته في تصنيف صور سرطان الجلد. أولًا، تم تحميل النموذج النهائي المدرب وتطبيقه على الصور المستخلصة من مجموعة بيانات HAM10000، بعد إجراء المعالجة المسبقة اللازمة والتي شملت إعادة تشكيل الصور وتوحيد القيم باستخدام المتوسط والانحراف المعياري المحسوبين أثناء التدريب.

تم احتساب دقة النموذج من خلال مقارنة الفئات المصنفة بالفئات الحقيقية، حيث تم حساب نسبة التصنيفات الصحيحة إلى إجمالي عدد الصور. بالإضافة إلى ذلك وكما أشرنا سابقًا، تم عرض مصفوفة الالتباس Confusion Matrix واستخراج تقرير التصنيف Classification Report والذي وفّر مقاييس تفصيلية لكل صنف. ساعدت هذه المقاييس في تقديم رؤية تحليلية لأداء النموذج، وتحديد نقاط القوة والضعف في التنبؤ بمختلف أصناف سرطان الجلد.

النتائج والمناقشة

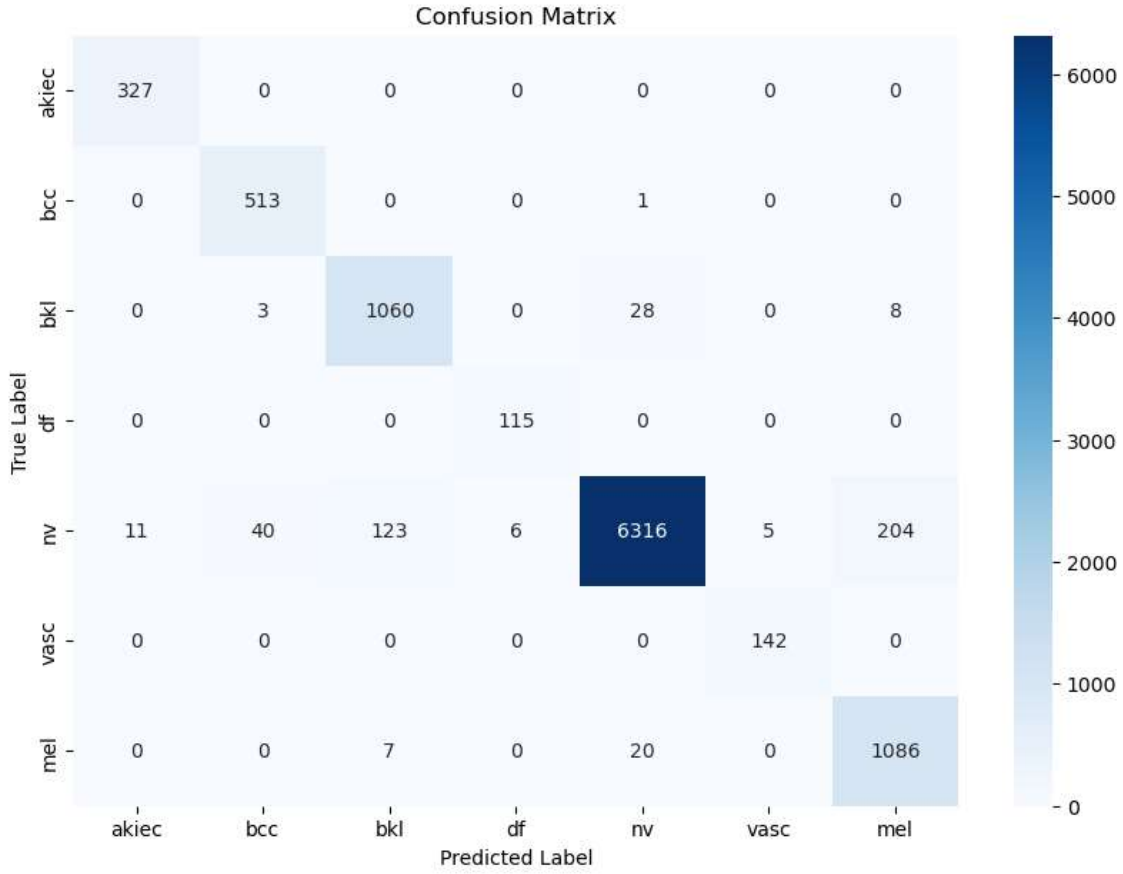
تقييم النموذج على مجموعة الاختبار

بعد الانتهاء من التدريب، تم تقييم مجموعة الاختبار التي بلغ حجمها 10,015 صورة، وتمكن النموذج من تصنيف 9559 منها بشكل صحيح، بينما أخطأ في تصنيف 456 صورة فقط، أي أن الدقة الإجمالية Accuracy لنموذج CNN المستخدم بلغت **95.45%**.

يوضح تقرير التصنيف Classification Report شكل (7) أن النموذج حقق اداءً مرتفعًا في جميع الفئات السبعة، حيث تراوحت الدقة لكل فئة Precision بين 84% و99%. فعلى سبيل المثال، فعلى سبيل المثال، بلغت الدقة في فئة *'akiec'* ≈ 97%، وفي فئة *'nv'* ≈ 99%، وفي فئة *'vasc'* ≈97%، بينما كانت أقل قيمة عند فئة *'mel'* ≈ 84%. بلغ المتوسط المرجّح لكل من Precision و F1-score حوالي **96%**، ما يعكس قدرة قوية للنموذج على التمييز بين الفئات المختلفة لسرطان الجلد. كما أظهرت مصفوفة الالتباس Confusion Matrix شكل (8) أن النموذج أدى بشكل ممتاز في معظم الفئات، مع ملاحظة وجود تداخل محدود بين فئتي nv وmel، وهو أمر متوقع نظرًا للتشابه البصري بين هذين النوعين في بعض الحالات السريرية.

	precision	recall	f1-score	support
akiec	0.97	1.00	0.98	327
bcc	0.92	1.00	0.96	514
bkl	0.89	0.96	0.93	1099
df	0.95	1.00	0.97	115
nv	0.99	0.94	0.97	6705
vasc	0.97	1.00	0.98	142
mel	0.84	0.98	0.90	1113
accuracy			0.95	10015
macro avg	0.93	0.98	0.96	10015
weighted avg	0.96	0.95	0.96	10015

الشكل 7: تقرير التصنيف Classification Report.

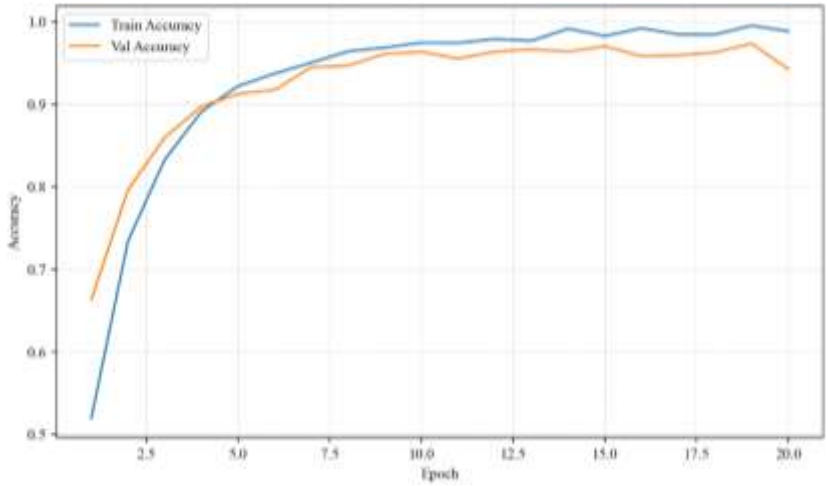


الشكل 8: مصفوفة الالتباس Confusion Matrix.

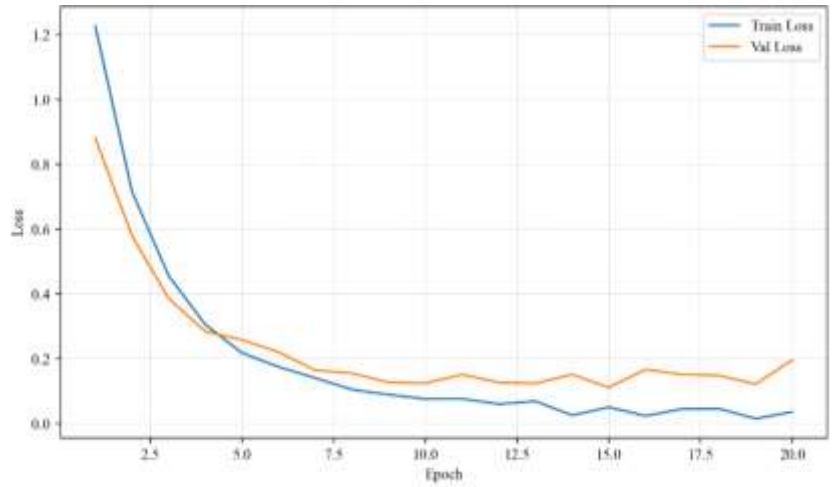
تحليل منحنيات التعلم

تُظهر المنحنيات في الشكلين (9) و(10) سلوك دقّة التدريب والتحقق، وكذلك الخسارة عبر 20 حقبة .يوضح الشكل (9) تحسنًا سريعًا في المراحل الأولى من التدريب، ثم استقرارًا وتقاربًا ملحوظًا بين دقّة التدريب والتحقق في معظم الحقَب، مما يشير إلى عدم وجود إفراطٍ واضح في التعلّم Overfitting على مستوى الدقّة. مع ذلك، يظهر الشكل (10) أن خسارة التدريب Training loss انخفضت إلى قيم أقل من خسارة التحقق Validation loss في بعض الحقَب، وهو أمر شائع في النماذج العميقة ويعكس زيادة ثقة النموذج في بيانات التدريب مقارنة ببيانات التحقق، لكنه لم يصل إلى درجة الإفراط الحاد.

تُبرز هذه النتائج دور استراتيجية Random Oversampling في تحسين التوازن بين الفئات النادرة والأكثر شيوعًا، ما ساهم في رفع الأداء على بيانات التحقق إلى ≈97%. ومع ذلك، لا يمكن استبعاد أن الاعتماد على Random Oversampling فقط قد يؤدي نظريًا إلى تضخيم الأداء بسبب تكرار عينات مشابهة، وهو ما قد يرفع احتمال حدوث Overfitting جزئي .إلا أن استقرار منحنيات التعلّم ونتائج الاختبار المرتفعة يوحي بأن هذا التأثير كان محدودًا في هذه الدراسة.



الشكل 9 :منحنى تغيّر الدقّة (Training vs Validation Accuracy) عبر الحقَب.



الشكل 10: منحنى تغيّر الخسارة (Training vs Validation Loss) عبر الحقَب

مناقشة النتائج

تشير هذه النتائج إلى أن النموذج المقترح حقق أداءً قويًا رغم التحديات المرتبطة بتعدد الفئات والتشابه البصري الكبير بين بعضها (مثل nv و mel) يمكن تفسير هذه الدقّة المرتفعة بعدة عوامل:

1. **كفاءة بنية الشبكة العصبية الالتفافية (CNN):** لقدرتها على استخلاص السمات المميّزة من الصور الطبية.
2. **استخدام تقنية Random Oversampling** :الذي عالج مشكلة عدم توازن البيانات ومنع انحياز النموذج للفئات الشائعة.
3. **جودة بيانات التدريب** :الصور كانت واضحة وممثلة بشكل جيد لمعظم الحالات.
4. **آلية التحقق الداخلي** :التي ساعدت على ضبط النموذج وحفظ أفضل نسخة أثناء التدريب.

وتُعتبر جميع هذه العوامل مؤشرًا قويًا على قدرة النموذج على التعميم، مما يجعله مناسبًا كخطوة نحو دعم القرار الطبي في تشخيص سرطان الجلد.

مقارنة بين نتائج العمل الحالي والدراسات السابقة

يوضح **الجدول** (4)المقارنة التفصيلية بين العمل الحالي والدراسات السابقة من حيث مجموعة البيانات، نوع النموذج، عدد الفئات المصنفة، تقنيات المعالجة المسبقة، والدقة المحققة.

لقد حقق النموذج المستخدم في هذا البحث دقّة عالية بلغت **95.45**% في تصنيف سبع فئات من آفات الجلد، متفوقًا بوضوح على الدراسات التي تناولت التصنيف متعددة الفئات: فقد بلغت دقّة الدراسة [15]**77.91%** فقط، بينما سجّلت الدراسة [4] في أفضل تشكيل Ensemble دقّة **82.37%**هذا يؤكد كفاءة النموذج واستقراره رغم استخدامه لبنية خفيفة ومعالجة مسبقة مقتصرة على Random Oversampling وإعادة تحجيم وتوحيد الصور، دون الحاجة لتعقيدات Augmentation أو Ensembles المتعددة. بالتالي، يثبت هذا العمل أن البساطة لا تقف عائقًا أمام الأداء العالي في مهام التصنيف متعددة الفئات، بل قد تكون عاملًا رئيسيًا في تقليل زمن التدريب والاستدلال وزيادة سهولة النشر العملي.

التوصيات

انطلاقاً من تحليل الأداء ونتائج الاختبارات التي أُجريت على مجموعة بيانات HAM10000وحرصًا على تعزيز دقة النموذج وموثوقيته وتوسيع إمكانياته التطبيقية في المجال السريري، تم صياغة التوصيات التالية:

1. دمج مجموعات بيانات من مصادر متعددة تغطي تنوع درجات لون البشرة وأنماط الجلد المختلفة (مثلISIC Archive و Derm7pt) لضمان قدرة النموذج على العمل بدقة مع جميع أنواع البشرة.

الجدول 4 : مقارنة العمل الحالي بالدراسات البحثية السابقة.

المرجع	مجموعة البيانات	نسبة التقسيم	نوع التصنيف	المعالجة المسبقة وتقنيات التعزيز	النماذج المستخدمة	دقة الاختبار%
[14]	ISIC 2018	80% تدريب 20% اختبار	ثنائي (حميد/خبيث)	ESRGAN، تدوير، انعكاس، تغيير حجم، تسوية	InceptionV3, ResNet50, Inception-ResNet	InceptionV3: 85.7
[15]	HAM 10000	67.5% تدريب 25% اختبار 7.5% تحقق	متعدد فئات (7 أنواع)	استبدال قيم مفقودة، تصغير أبعاد، تدوير، تكبير، انعكاس أفقي/عمودي	CNN	77.91
[16]	ISIC(ثنائي: حميد مقابل خبيث)	80% تدريب 20% اختبار	ثنائي (حميد/خبيث)	ISR (Image Super-Resolution)	VGG16, ResNet, InceptionV3	InceptionV3: 91.26
[17]	ISIC (حميدة، خبيثة)	80% تدريب 20% اختبار	ثنائي (حميدة/خبيثة)	تغيير حجم، تدوير، إزاحة أفقية/عمودية، تكبير عشوائي، تسوية لونية	NASNet Mobile & NASNet Large	NASNet Mobile:85.62 NASNet Large:83.98
[8]	Skin Cancer: Malignment vs Banging	80% تدريب 20% اختبار	ثنائي (حميد/خبيث)	إزالة تكرارات، توحيد الحجم، تدوير، إزاحة، قص، تكبير	DenseNet169	89.7
[4]	ISIC 2019	80% تدريب 10% اختبار 10% تحقق	متعدد فئات (7 أنواع)	قص/تقليم، تعزيز البيانات Data Augmentation	Ensemble (InceptionV3, ResNet50, VGG16, Xception…)	VGG16+ Xception: 82.37
العمل الحالي	HAM 10000	64% تدريب 16% تحقق 20% اختبار	متعدد فئات (7 أنواع)	إعادة تشكيل الصور، توحيد، تقنية التجاوز العشوائي Random Oversampling لتوازن الفئات	CNN	95.45%

2. الاستفادة من نماذج مُدربة مسبقًا (مثل EfficientNet أو ResNet) وتطبيق أساليب تعزيز البيانات المتقدمة (Advanced Data Augmentation) لتعزيز تنوع الطبقات القليلة البيانات وتقليل الإفراط في التعلم.
3. معالجة الصور مسبقًا باستخدام خوارزمية إزالة الضوضاء وتقنيات مساواة التباين أو استخدام أساليب تحسين دقة الصورة لتعزيز تمييز حواف الأورام.
4. نشر النموذج على بيئة ويب خفيفة مع واجهة لرفع الصور وعرض النتائج، واختبار زمن الاستجابة على أجهزة محمولة مع ضمان تشفير البيانات والتوافق مع متطلبات الخصوصية.

الخاتمة

قدّم هذا البحث نموذجًا فعّالًا في التصنيف متعدد الفئات لصور سرطان الجلد بالاعتماد على بنية شبكية خفيفة الوزن وذات كفاءة عالية. تم تصميم النموذج باستخدام شبكات عصبية تلافيفية تسلسلية (CNN)، وحقق دقة اختبار بلغت 95.45٪ عند تصنيف سبعة أنواع من آفات الجلد، ما يدل على قدرته العالية على التمييز بين الفئات المختلفة. يُبرز هذا النموذج إمكانيات الحلول الذكية المبسطة في دعم التشخيص الطبي، خاصة في البيئات ذات الإمكانيات المحدودة. كما يؤكد على أهمية تحقيق التوازن بين الأداء والدقة من جهة، ومتطلبات الموارد الحسابية من جهة أخرى، مما يجعله مناسبًا للتطبيق العملي في نظم التشخيص الطبي المعتمد على صور.

Author Contributions: "Fatima Ahmed and Al-Zahraa Othman: Conceptualization, methodology; Fatima Ahmed: writing—original, review and editing; Ali Ukasha: supervision. All authors have read and agreed to the published version of the manuscript."
Funding: "This research received no external funding".
Data Availability Statement: "The data are available at request."
Conflicts of Interest: "The authors declare that they have no conflict."

References

[1] I. Pacal, B. Ozdemir, J. Zeynalov, H. Gasimov, and N. Pacal. "A novel CNN-ViT-based deep learning model for early skin cancer diagnosis," *Biomedical Signal Processing and Control*, vol. 104, pp. 107627, 2025. https://doi.org/10.1016/j.bspc.2025.107627.

[2] M. Dildar, et al. "Skin cancer detection: a review using deep learning techniques," *International journal of environmental research and public health*, vol. 18, no. 10, pp. 5479, 2021. https://doi.org/10.3390/ijerph18105479.

[3] T. Kumar, and I. Himanshu. "Artificial intelligence based real-time skin cancer detection," in 2023 15th International Conference on Computer and Automation Engineering (ICCAE), 2023, pp. 215-219. https://doi.org/10.1109/ICCAE56788.2023.10111099.

[4] K. Natarajan, D. Kumar, G. Murali, C. Madhuvappan, S. Kannan, and M. Mathesh. "Enhancing Skin Cancer Diagnosis with Ensemble Deep Convolutional Neural Networks: A Multi-Model Optimization Approach," in 2024 Second International Conference on Inventive Computing and Informatics (ICICI), 2024, pp. 412-416. https://doi.org/10.1109/ICICI62254.2024.00074.

[5] A. Shahin, A. Kamal, and M. Elattar. "Deep ensemble learning for skin lesion classification from dermoscopic images," in 2018 9th Cairo International Biomedical Engineering Conference (CIBEC), 2018, pp. 150-153. https://doi.org/10.1109/CIBEC.2018.8641815.

[6] E. Houssein, D. Abdelkareem, G. Hu, M. Hameed, I. Ibrahim, and M. Younan. "An effective multiclass skin cancer classification approach based on deep convolutional neural network," *Cluster Computing*, vol. 27, no. 9, pp. 12799-12819, 2024. . https://doi.org/10.1007/s10586-024-04540-1.

[7] T. Mazhar, et al. "The role of machine learning and deep learning approaches for the detection of skin cancer," in Healthcare, 2023, pp. 415. https://doi.org/10.3390/healthcare11030415.

[8] R. Pathania, and P. Behki, "Skin cancer detection using deep learning," in 2024 Sixth International Conference on Computational Intelligence and Communication Technologies (CCICT), 2024, pp. 568-575. https://doi.org/10.1109/CCICT62777.2024.00095.

[9] "Dataset: Skin Cancer MNIST: HAM10000," https://www.kaggle.com/datasets/kmader/skin-cancer-mnist-ham10000. [Accessed: March 2,2025].

[10] F. Mishri, and R. Ali. "Detection and Classification of Brain Tumours from MRI Images Using VGG16 Deep Learning Algorithm," *Wadi Alshatti University Journal of Pure and Applied Sciences*, pp. 43-50, 2025. https://doi.org/10.63318/.

[11] F. Murat, O. Yildirim, M. Talo, U. Baloglu, Y. Demir, and U. Acharya. "Application of deep learning techniques for heartbeats detection using ECG signals-analysis and review," *Computers in biology and medicine*, vol. 120, pp. 103726, 2020. https://doi.org/10.1016/j.compbiomed.2020.103726.

[12] Y. Wu, B. Chen, A. Zeng, D. Pan, R. Wang, and S. Zhao. "Skin cancer classification with deep learning: a systematic review," *Frontiers in Oncology*, vol. 12, pp. 893972, 2022. https://doi.org/10.3389/fonc.2022.893972.

[13] T. Mendonça, P. Ferreira, J. Marques, A. Marcal, and J. Rozeira. "PH 2-A dermoscopic image database for research and benchmarking," in 2013 35th annual international conference of the IEEE engineering in medicine and biology society (EMBC), 2013, pp. 5437-5440. https://doi.org/10.1109/EMBC.2013.6610779.

[14] W. Gouda, N. Sama, G. Al-Waakid, M. Humayun, and N. Jhanjhi. "Detection of skin cancer based on skin lesion images using deep learning," in Healthcare, 2022, pp. 1183. https://doi.org/10.3390/healthcare10071183.

[15] Y. Shobha, K. Prasad, and A. SG. "Multiclass Classification of Skin Cancer using Convolutional Neural Network," in 2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon), 2022, pp. 1-5. https://doi.org/10.1109/MysuruCon55714.2022.9972486.

[16] A. Lembhe, P. Motarwar, R. Patil, and S. Elias,. "Enhancement in skin cancer detection using image super resolution and convolutional neural network," *Procedia Computer Science*, vol. 218, pp. 164-173, 2023. https://doi.org/10.1016/j.procs.2022.12.412.

[17] M. A. Rahman, E. Bazgir, S. Hossain, and M. Maniruzzaman. "Skin cancer classification using NASNet," *International Journal of Science and Research Archive*, vol. 11, no. 1, pp. 775-785, 2024. https://doi.org/10.30574/ijsra.2024.11.1.0106.

[18] P. Tschandl, C. Rosendahl, and H. Kittler. "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Scientific data*, vol. 5, no. 1, pp. 1-9, 2018. https://doi.org/10.7910/DVN/DBW86T.

[19] K. O'shea, and R. Nash. "An introduction to convolutional neural networks," *arXiv preprint arXiv:1511.08458*, 2015. https://doi.org/10.48550/arXiv.1511.08458.

[20] A. Khan, A. Sohail, U. Zahoora, and A. Qureshi. “A survey of the recent architectures of deep convolutional neural networks,” *Artificial intelligence review*, vol. 53, pp. 5455-5516, 2020. https://doi.org/10.1007/s10462-020-09825-6.

[21] L. Chen, S. Li, Q. Bai, J. Yang, S. Jiang, and Y. Miao. “Review of image classification algorithms based on convolutional neural networks,” *Remote Sensing*, vol. 13, no. 22, pp. 4712, 2021. https://doi.org/10.3390/rs13224712.

[22] R. Bogne Tchema, A. Polycarpou, and M. Nestoros. “Skin cancer classification using machine learning,” *Multimedia Tools and Applications*, vol. 84, no. 6, pp. 3239-3256, 2025. https://doi.org/10.1007/s11042-025-20595-7.

[23] H. Asim. "Skin Cancer Detector," 2024 .[online]. Available: https://github.com/HassanAsim/Skin-Cancer-Detector/blob/main/Skin_Cancer_Detector.ipynb. [Accessed: March 2, 2025].

[24] K. Sujon, R. Hassan, Z. Towshi, M. Othman, M. Samad, and K. Choi. “When to Use Standardization and Normalization: Empirical Evidence from Machine Learning Models and XAI,” *IEEE Access*, 2024. https://doi.org/10.1109/ACCESS.2024.3462434.

[25] M. Buda, A. Maki, and M. Mazurowski. “A systematic study of the class imbalance problem in convolutional neural networks,” *Neural networks*, vol. 106, pp. 249-259, 2018. https://doi.org/10.1016/j.neunet.2018.07.011.

[26] M. Omar, and A. Ukasha. “Using Deep Learning and Particle Swarm Optimization Algorithm to Recognize Fatty Liver in Compressed Images,” *Wadi Alshatti University Journal of Pure and Applied Sciences*, pp. 119-126, 2025. https://doi.org/10.63318/waujpasv3i1_18 .