DYNAMIC DETECTION OF SURFACE CRACKS USING DEEP LEARNING AND COMPUTER VISION

Major Project –II (CV499) Report
submitted in partial fulfilment of the requirement for the degree of

BACHELOR OF TECHNOLOGY
in
Civil Engineering

by

Shashank Pandey (201CV150)
Tuhin Anand (201CV155)
Parsapu Kamal Raj (201CV170)

DEPARTMENT OF CIVIL ENGINEERING
NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA
MANGALORE -575 025

APRIL 2024

DYNAMIC DETECTION OF SURFACE CRACKS USING DEEP LEARNING AND COMPUTER VISION

Major Project –II (CV499) Report
submitted in partial fulfilment of the requirement for the degree of

BACHELOR OF TECHNOLOGY
in
Civil Engineering

by

Shashank Pandey (201CV150)
Tuhin Anand (201CV155)
Parsapu Kamal Raj (201CV170)

under the guidance of
Dr. Palanisamy T.

DEPARTMENT OF CIVIL ENGINEERING
NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA
MANGALORE -575 025

APRIL 2024

DECLARATION

We hereby declare that the report of the Major project-II (CV499) entitled **DYNAMIC DETECTION OF SURFACE CRACKS USING DEEP LEARNING AND COMPUTER VISION,** which is being submitted to **National Institute of Technology Karnataka, Surathkal** in partial fulfilment of requirements for the award of the degree of **Bachelor of Technology in Civil Engineering** is a bonafide record of the work carried out by us. The material contained in this report has not been submitted to any university or Institution for the award of any degree

1. Shashank Pandey (201CV150)
2. Tuhin Anand (201CV155)
3. Parsapu Kamal Raj (201CV170)

Place: NITK, Surathkal
Date: April 2024

CERTIFICATE

This is to certify that this report entitled **DYNAMIC DETECTION OF SURFACE CRACKS USING DEEP LEARNING AND COMPUTER VISION**, submitted by,

1. Shashank Pandey (201CV150)
2. Tuhin Anand (201CV155)
3. Parsapu Kamal Raj (201CV170)

as the record of the work carried out by them in the Department of Civil Engineering, is accepted as the B.Tech Major Project –II Report (CV499) submission in partial fulfilment of the requirement for the award degree of **Bachelor of Technology in Civil Engineering.**

Dr. Palanisamt T.
Associate Professor
Department of Civil Engineering

Chairperson – DUGC
Department of Civil Engineering

ACKNOWLEGEMENTS

This work is not just a honest effort by a group of three students but also has been a combined effort of everyone who supported us throughout this project. First our respected guide and Associate Professor Dr. Palanisamy T. has been instrumental to us throughout this project with his continuous suggestions, discussions, evaluations and proper guidance.

Our HOD, Prof. Subhash C Yaragal has supported us throughout this project and has been a constant source of motivation for us.

Last but not the least our friends who helped us out whenever we were stuck with something in the project. With all this guidance and support, we were able to complete this project.

TABLE OF CONTENTS

Table of Contents

TABLE OF FIGURES

CHAPTER 1

INTRODUCTION

General:

In civil engineering, a crack refers to a visible fissure or separation in a structure, such as a building or pavement, often caused by factors like stress, settlement, or material deterioration. Cracks can impact the integrity of the structure and may require evaluation and repair to ensure safety and longevity.

Figure 1 CRACK IN STRUCTURE

In other words cracks are physical manifestations of the response of a structure to various forces and environmental conditions. These forces can include factors like load-bearing stress, temperature fluctuations, settlement, or the natural aging of materials. When a structure experiences stress beyond its capacity or encounters conditions that compromise its integrity, it may develop cracks.

Cracks can vary in size, shape, and orientation, and they are classified based on factors such as their width, depth, and location within the structure. Common types of cracks include hairline cracks, settlement cracks, and structural cracks. Understanding the type and cause of a crack is crucial for engineers to determine the appropriate remedial measures.

Occurrence of various crack patterns in the building during construction, after completion when it is subjected to super imposed load or during the service life, is a common phenomenon.

Therefore, effective crack management is important. Effective crack management involves thorough inspection, monitoring, and, if necessary, implementing repairs to prevent further deterioration. Repair methods may include filling the crack with specialized materials, reinforcing the structure, or addressing the underlying issues that led to the crack formation.

Impact of cracks on structure:

Cracks in concrete and steel structures can have significant impacts on the structural integrity, durability, and safety of the infrastructure. The consequences of cracks depend on various factors, including the type of structure, the size and location of the cracks, and the underlying causes.

Here are some general impacts of cracks on concrete and steel structures:

1. Structural Integrity:

Cracks can compromise the structural integrity of a building or bridge, reducing its load-carrying capacity.

In reinforced concrete structures, cracks may lead to the corrosion of embedded steel reinforcement, further weakening the structure.

2. Durability:

Cracks provide pathways for water, chemicals, and aggressive environmental agents to penetrate the structure. This can accelerate the deterioration of the material and decrease the overall durability.

Freeze-thaw cycles can exacerbate cracks, leading to spalling and further damage.

3. Corrosion of Steel Reinforcement:

Cracks in concrete expose embedded steel reinforcement to moisture and oxygen, leading to corrosion. Corrosion can result in the expansion of steel, causing additional cracking and reducing the strength of the structure.

4. Serviceability Issues:

Cracks can affect the serviceability of a structure by causing water leakage, air infiltration, and aesthetic concerns.

In buildings, cracks may lead to difficulty in maintaining a controlled indoor environment.

5. Safety Concerns:

Large or critical cracks may pose safety hazards, especially in structures like bridges and high-rise buildings.

Sudden failure or collapse can occur if cracks are not addressed promptly and appropriately.

6. Repair and Maintenance Costs:

Repairing cracks can be costly and time-consuming. Ignoring cracks or implementing inadequate repairs may result in recurring maintenance expenses.

Early detection and proper repair methods can help minimize long-term costs.

7. Impact on Functionality:

Cracks can affect the functionality of structural elements such as beams, columns, and slabs, potentially leading to the need for alterations or additional support.

8. Environmental Impact:

The environmental impact of cracks includes the potential for material wastage, increased energy consumption during repair processes, and the release of pollutants from repair materials.

Types of cracks:

Cracks in buildings could be broadly classified as structural and non-structural cracks.

1. Structural Cracks: These occur due to incorrect design, faulty construction or overloading and these may endanger the safety of a building.

Example - Extensive cracking of an RCC beam.

Figure 2 STRUCTURAL CRACKS

2. Non-structural Cracks: These are mostly due to internally induced stresses in buildings materials and do not endanger safety of a building but may look unsightly, or may create an impression of faulty work or may give a feeling of instability.

Figure 3 NON-STRUCTURAL CRACKS

In some situations due to penetration of moisture through them non-structural cracks may spoil the internal finishes thus adding to the cost of maintenance, or corrode the reinforcement, thereby adversely affecting the stability of the Structure in long run. e.g. Vertical crack in a long compound wall due to shrinkage or thermal movement.

Apart from these, among non-structural type of cracks there are numerous different categories of cracks. Some of them are:

Hairline Cracks:

Extremely narrow cracks that are often cosmetic and do not significantly affect the structural integrity, caused by shrinkage during the curing process of concrete or minor settlement.

Figure 4 HAIRLINE CRACK

Crazing:

Fine, shallow cracks on the surface of concrete, caused by the rapid drying of the surface during the curing process.

Figure 5 CRAZING

Expansion Cracks:

Cracks, which occur due to the expansion of materials such as concrete, when exposed to high temperatures. Common in structures subject to temperature variations.

Figure 6 EXPANSION CRACKS

Settlement Cracks:

Cracks which occur due to the settling of a structure or its foundation. Uneven settlement can lead to differential movement and cracking.

Figure 7 SETTLEMENT CRACKS

Freeze-Thaw Cracks:

Cracks caused by the repeated freeze-thaw cycles in cold climates. Water penetrates the structure, freezes, and expands, leading to cracking.

Figure 8 FREEZE-THAW CRACKS

Corrosion-Induced Cracks:

Cracks resulting from the corrosion of embedded steel reinforcement. Corrosion products can cause the steel to expand, leading to cracking.

Figure 9 CORROSION-INDUCED CRACKS

Causes of cracks:

Cracks in structures can result from a variety of causes, and understanding the underlying factors is essential for effective repair and prevention. Here are some common causes of cracks in structures:

Shrinkage:

Concrete undergoes shrinkage as it cures and loses water. This can lead to the development of hairline cracks, especially in large, unreinforced concrete slabs.

Settlement:

Uneven settlement of the foundation can cause differential movement in the structure, leading to cracks. This can occur due to poor soil conditions, inadequate foundation design, or changes in groundwater levels.

Overloading:

Excessive loads beyond the design capacity of a structure can cause deformation and cracking. This may result from additional live loads, improper use of the structure, or changes in occupancy.

Temperature Changes:

Thermal expansion and contraction due to temperature variations can result in cracks, especially in materials like concrete. Rapid temperature changes can lead to thermal stresses.

Corrosion of Reinforcement:

Corrosion of steel reinforcement within concrete can cause it to expand, leading to cracking and spalling. Corrosion is often accelerated by exposure to aggressive environmental conditions or the use of deicing salts.

Poor Construction Practices:

Inadequate construction techniques, such as improper mix proportions, inadequate curing, or insufficient compaction, can contribute to the development of cracks.

Water Intrusion:

Water infiltration into the structure can weaken the concrete and cause expansion and contraction, leading to cracking. Freeze-thaw cycles can exacerbate this issue.

Chemical Attack:

Exposure to aggressive chemicals, such as sulfates, chlorides, or acids, can deteriorate the concrete matrix, leading to the development of cracks.

Differential Movement:

Variations in movement between different parts of a structure can result in cracks. This can occur due to settlement, differential loading, or inadequate provision of expansion joints.

Earthquakes:

Seismic activity can induce dynamic forces on structures, leading to cracking. Proper seismic design and reinforcement are essential to minimize earthquake-induced damage.

Design Flaws:

Errors or omissions in the structural design can lead to stress concentrations and cracking. Inadequate detailing or improper consideration of loads can contribute to structural failures.

Tree Roots:

Tree roots can exert significant pressure on foundations, leading to movement and cracking. This is particularly common in older structures with mature trees nearby.

Construction Settlement:

Settlement that occurs during or immediately after construction can result in cracks as the structure adjusts to its final position.

Identifying the specific cause of cracks is crucial for implementing appropriate remedial measures. Regular inspections, proper construction practices, and adherence to design and maintenance guidelines can help mitigate the risk of cracks in structures.

Measures for prevention of cracks:

Preventing cracks in structures involves a combination of proper design, construction practices, and maintenance. Here are some measures to help prevent the formation of cracks in buildings and other structures:

Proper Design:

Ensure that the structure is designed by qualified engineers with consideration given to the type of structure, loads, and environmental conditions.

Include appropriate structural elements, such as expansion joints and control joints, to accommodate movements and reduce stress.

Quality Construction Materials:

Use high-quality materials that meet or exceed industry standards for construction. Follow proper mix proportions for concrete to ensure adequate strength and durability.

Controlled Curing:

Implement proper curing procedures for concrete to minimize shrinkage and enhance the strength of the material.

Protect newly placed concrete from rapid drying and extreme temperature variations during the curing period.

Construction Practices:

Adhere to proper construction techniques and methodologies.

Ensure accurate placement and consolidation of concrete to reduce the likelihood of voids and weak spots.

Foundation Design:

Conduct thorough site investigations to assess soil conditions and design foundations that account for potential settlement.

Use appropriate foundation types based on soil characteristics and loads.

Proper Reinforcement:

Adequately reinforce concrete structures to resist tensile forces and prevent cracking.

Use corrosion-resistant reinforcement to minimize the risk of steel corrosion-induced cracking.

Expansion Joints:

Install expansion joints in areas where large temperature-induced movements or settlement differentials are anticipated.

Expansion joints allow controlled movement and help prevent uncontrolled cracking.

Moisture Control:

Implement proper waterproofing measures to protect structures from water infiltration.

Ensure adequate drainage around foundations to prevent soil-related issues.

Regular Inspections:

Conduct regular inspections to identify potential issues before they escalate.

Monitor cracks and movement, and address them promptly to prevent further damage.

Proper Load Distribution:

Design and construct structures to distribute loads evenly to prevent localized stress concentrations. Avoid overloading structures beyond their design capacity.

Seismic Design:

In earthquake-prone regions, follow seismic design codes and guidelines to ensure structures can withstand seismic forces without significant damage.

Maintenance and Repairs:

Implement a proactive maintenance program to address issues before they become severe.

Perform timely repairs to address cracks and other signs of distress.

Soil Stabilization:

Implement measures to stabilize soil conditions and reduce the risk of settlement-related cracks. This may include proper compaction, soil improvement, or the use of geotechnical solutions.

Education and Training:

Ensure that construction personnel are well-trained in proper construction practices and are aware of factors that can lead to cracking.

By incorporating these preventive measures into the design, construction, and maintenance processes, it is possible to reduce the likelihood of cracks in structures and enhance their overall durability and performance.

CHAPTER 2

LITERATURE REVIEW

Yiyang et al. have proposed a crack detection algorithm based on digital image processing technology. By pre- processing, image segmentation and feature extraction [4], they have obtained the information about the crack image. In, Threshold method of segmentation was used after the smoothening of the accepted input image. To judge their image, they have calculated the area and perimeter of the roundness index. Then by the comparison, they have evaluated the presence of the crack in the image.

Alam et al. have proposed a detection technique by the combination of the digital image correlation and acoustic emission. The former method gives a very precise measurement of surface displacements, thus crack openings and crack spacing were determined. In order to complement that method and to investigate damage mechanisms, acoustic emissions resulting from internal damage were also analyzed. A manual grouping method (similar to K-means method) was used to identify different classes of AE energy released from the Beams of three different sizes. In their methodology, they have used three different beam proportionalities for the effectiveness of the output.

Talab et al. have presented a new approach in image processing for detecting cracks in images of concrete structures. Here the methodology involves three steps: First; change the image to a gray image using the edge of the image and then use Sobel's method to develop an image using Sobel's filter for detecting cracks. Then by using suitable threshold binary image of the pixel they are categorized into the foreground and the background image. Once the images are categorized, Sobel's filtering was used for the elimination of residual noise. After the vast filtering procedure of the image, cracks were detected using the otsu's method. They have replaced the sobel filter with the multiple median filtering in certain cases.

Yamaguchi et al. have developed a percolation-based crack detection technique. They have obtained their less computation time by the adaptation of the termination and skip add procedures. They have a high-speed percolation algorithm which will make use of the neighboring pixels based upon the circularity of the pixel needs. The template matching technique was the key to their proposal of percolation because matching in the percolation images was easy to analyze.

Yang et al. have proposed an image analysis method to capture thin cracks and minimize the requirement for pen marking in reinforced concrete structural tests. They have used the studies like crack depth prediction, change in detection without image registration, crack pattern recognition based on artificial neural networks, applications to micro-cracks of rocks, and efficient sub-pixel width measurement. Stereo triangulation method was the adopted technique based on cylinder formula approximation and image rectification. Once they have the rectified output, the surface of the observed regions can be unfolded and presented in a plane image for following displacement and deformation analysis. From which the crack detection was analyzed.

Sinha et al. have investigated the cracks by using the two-step approach. They have developed a statistical filter design for the crack detection. After the filtering, they have got to the two-step approach at which the crack feature extraction was done locally at the first step of the pre-processing and then they have fused the images. The second step is to define the crack among the image segment by the process of cleaning and linking. They have overcome their previous work disadvantage where the morphological approach was used.

Wang et al. have proposed a system for the image based crack detection and to characterize the crack based upon their effectiveness. They have categorized the present image based crack detection into four categories. They are an integrated algorithm, morphological approach, percolation approach and practical technique. A shading correction was done using integrated algorithm. The unclear crack prediction was detected using percolation method. The crack detection was done using morphological approach for the micro crack detection with the practical method providing high-performance feature extraction.

Xu et al. have proposed a system on infrared thermal image processing frame work based on super pixel to detect the crack. The segmentation was done based on the Fuzzy c-means clustering. The generation of the super pixels has been done because of its adherence to crack boundaries. The super pixels were selected from the raw gray image as well as high pass filtered image.

Lee et al. have designed a system for particle crack detection. They used the nearest neighbor and two-point correlation methods for the estimation of the second order microstructural descriptors. Based on the probability function of their corresponding location the crack features were found out. The edge effect was eliminated by the nearest neighbor estimate from the high-resolution montages.

Oliveira et al have designed a system for the automatic crack detection. Here the crack detection was based on the sample paradigm. In the sample paradigm, a subset of the available image database was automatically selected and used for unsupervised training of the system images. They have characterized operations based on the classification of the non-overlapping image blocks. Then based on the crack block based detection, the width of the crack was estimated.

Conclusion from mentioned research work and papers:

The literature reviews presented here highlight various methodologies and techniques employed by different researchers for the detection of cracks in structures, particularly focusing on digital image processing technology. Each study introduces a unique approach, combining different methods and technologies to achieve accurate and efficient crack detection. Here are some key conclusions drawn from the literature reviews:

Diverse Approaches:

The reviewed studies employ diverse approaches, including digital image processing, acoustic emission, percolation-based techniques, infrared thermal imaging, and statistical filtering. This diversity reflects the multifaceted nature of crack detection challenges and the need for versatile methodologies.

Integration of Technologies:

Several studies combine multiple technologies to enhance the accuracy and reliability of crack detection. For example, Alam et al. combine digital image correlation with acoustic emission analysis, and Xu et al. integrate infrared thermal imaging with super pixel segmentation.

Pre-processing and Feature Extraction:

Common steps across the studies include pre-processing, image segmentation, and feature extraction. Techniques such as Sobel filtering, thresholding, percolation, and statistical filtering are utilized to extract meaningful information related to crack presence, size, and morphology.

Adaptation of Algorithms:

Researchers often adapt existing algorithms or develop novel ones to address specific challenges in crack detection. For example, Yiyang et al. use a threshold method of segmentation after image smoothening, while Talab et al. employ Sobel's filtering and Otsu's method for crack detection.

Efficiency and Computation Time:

Efficiency and computation time are critical considerations. Yamaguchi et al. emphasize high-speed percolation algorithms, and Yang et al. propose an image analysis method to capture thin cracks while minimizing the need for extensive processing.

Use of Machine Learning and deep learning:

While not explicitly mentioned in all studies, some research suggests the use of machine learning and deep learning techniques, such as artificial neural networks, for crack pattern recognition and feature extraction.

Application Specificity:

The studies often tailor their methodologies to specific applications, such as crack detection in concrete structures, particle crack detection, or automatic crack detection in various contexts. This underscores the importance of considering the unique challenges posed by different materials and structures.

Advancements in Image-Based Techniques:

Wang et al. categorize image-based crack detection into integrated algorithms, morphological approaches, percolation approaches, and practical techniques. This categorization reflects ongoing advancements in image-based methodologies for crack detection.

Automation and Unsupervised Training:

Oliveira et al. introduce an automated system for crack detection based on unsupervised training using a subset of the image database. This emphasizes the move towards automation and the use of machine learning for crack detection without the need for extensive manual input.

In conclusion, the literature reviews demonstrate a rich landscape of methodologies and technologies for crack detection in structures. The field is evolving with a focus on integration, efficiency, and adaptability to different applications. As technology continues to advance, there is a growing trend towards automation and the incorporation of machine learning techniques for more robust and accurate crack detection systems.

CHAPTER 3

MATERIALS AND METHODS

In this project we have utilised Deep learning, computer vision to a great extent. Although project is not perfect and certainly requires some more efficient modifications to be commercially or industrial used, it definitely can serve the testing purpose in a short run.

We have put our whole hearted efforts going into this project and have obtained some great results with the different models.

We classify the surface cracks using 3 different base models with transfer learning:

- ResNet v2
- VGG16
- Xception

Finally, we will display the activation maps using Grad-Cam.

The best feature is, it can be used in real time using phone camera and results can be simultaneously displayed or viewed on the laptop with high accuracy.

There is very slight time lag of 1 or 2 seconds when you are continuously moving the camera and suddenly focus on it on some specific crack or surface.

This issue could be addressed with a bit powerful computer system that has higher RAM.

In recent years, deep learning has emerged as a powerful tool in image analysis tasks, including object detection and classification. This project focuses on leveraging the capabilities of three state-of-the-art deep learning architectures: ResNet v2, VGG16, and Xception, for the detection of cracks in structural elements.

ResNet v2, known for its exceptional ability to handle the challenges of training very deep networks, is employed to capture intricate features within images. The introduction of skip connections in ResNet v2 facilitates the learning of residual functions, enabling more effective representation of complex crack patterns.

VGG16, with its straightforward architecture comprising multiple convolutional layers, is chosen for its simplicity and efficiency. The small (3x3) convolutional filters in VGG16 aid in learning hierarchical features, making it adept at capturing nuanced details in images of structural elements.

Xception, a model inspired by the Inception architecture, brings efficiency to the project through depthwise separable convolutions. By separating spatial and channel-wise operations, Xception optimizes parameter usage, making it well-suited for crack detection tasks where resource efficiency is crucial.

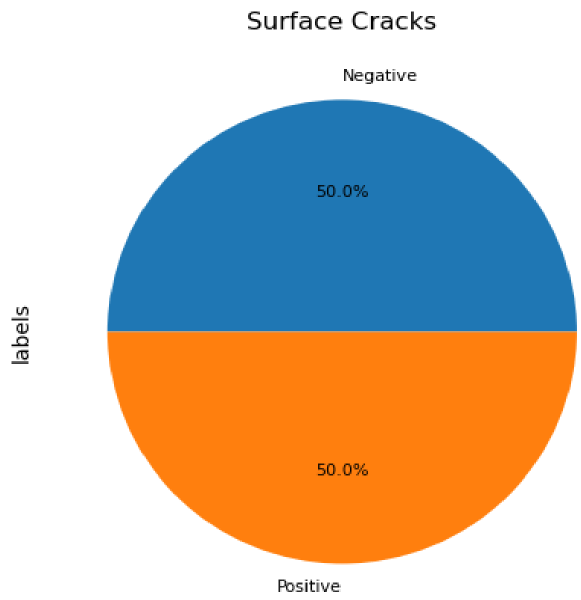


Figure 10 SURFACE CRACKS

In this project, we take a dataset comprising 40,000 images with positive and negative labels for cracks. The dataset includes 20,000 images labeled as crack-positive and 20,000 images labeled as crack-negative. This diverse dataset enables comprehensive training and evaluation of the models across various crack types and structural materials.

In this project, we aim to compare the performance of ResNet v2, VGG16, and Xception in the context of crack detection, considering factors such as accuracy, speed, and generalization to different types of cracks and structural materials. The insights gained from this study can contribute to the development of robust and efficient crack detection systems, enhancing the overall safety and resilience of critical infrastructure.

CHAPTER 4

RESULTS AND DISCUSSION

ResNet v2

ResNet v2, an evolution of the original ResNet architecture, represents a groundbreaking approach in deep neural networks. Introduced as a solution to the vanishing gradient problem, ResNet v2 incorporates a residual learning framework that utilizes identity connections to address the challenges associated with training extremely deep networks. Unlike its predecessor, ResNet v2 places the activation and batch normalization prior to weight layers, creating a pre-activation residual block. This strategic modification enables smoother gradient flow during backpropagation, enhancing the model's training efficiency. The architecture comprises bottleneck blocks, integrating 1x1, 3x3, and 1x1 convolutions, optimizing computational efficiency without compromising expressive power. The residual connections facilitate the training of exceedingly deep networks, with ResNet-152 being a notable example, achieving remarkable accuracy in image classification tasks. Furthermore, ResNet v2 introduces the concept of skip connections, allowing the direct flow of information between layers, thereby mitigating the risk of vanishing gradients. This innovative design not only enhances training but also contributes to improved model generalization and accuracy. Overall, ResNet v2 stands as a testament to the continual refinement of deep neural network architectures, emphasizing the importance of residual connections in facilitating the training of exceptionally deep networks with superior performance across various computer vision tasks.

Architecture

ResNet v2, or Residual Network version 2, is an extension and refinement of the original ResNet architecture. Its key features include a focus on pre-activation residual blocks and the integration of bottleneck design for computational efficiency.
The architecture is characterized by its deep structure, with ResNet-152 being a prominent variant

Pre-Activation Residual Blocks: Unlike the original ResNet, v2 introduces pre-activation residual blocks, where activation and batch normalization are applied before the convolutional layers. This modification facilitates smoother gradient flow during backpropagation, addressing the vanishing gradient problem encountered in training very deep networks.

Bottleneck Design: ResNet v2 adopts a bottleneck design in its residual blocks, which involves stacking three layers: 1x1, 3x3, and 1x1 convolutions. The 1x1 convolutions are employed to reduce and then restore the dimensions, optimizing computational efficiency. This design allows for

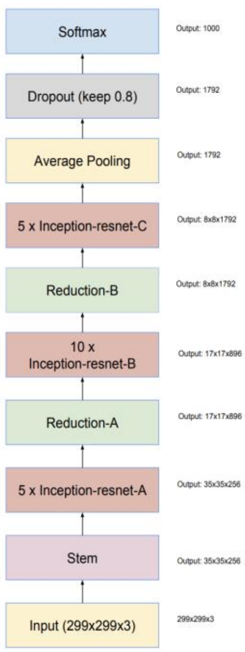


Figure 11 ARCHITECTURE FLOWCHART

the construction of deeper networks without a proportional increase in computational cost.

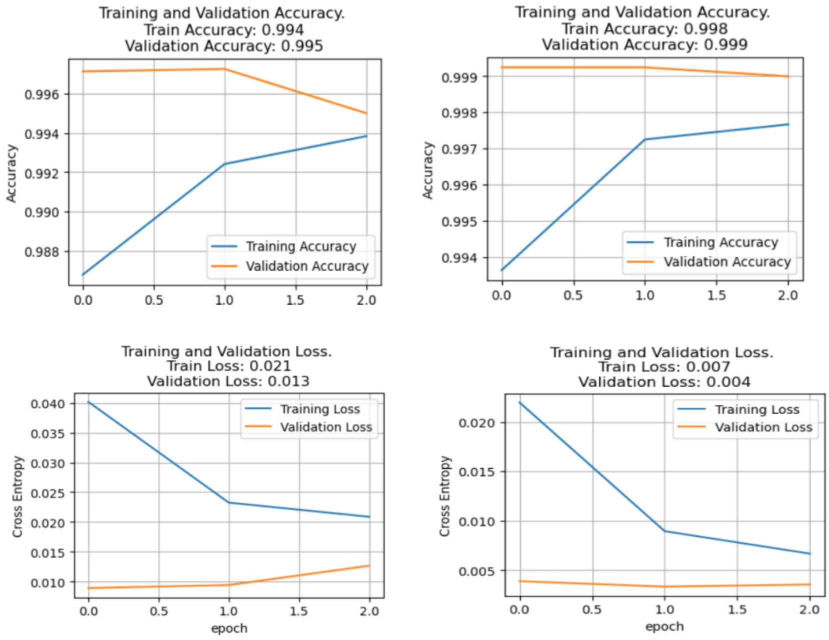
Skip Connections: ResNet v2 retains the skip connections from the original ResNet, enabling the direct flow of information between layers. These skip connections, also known as identity mappings, play a crucial role in mitigating the vanishing gradient problem. By facilitating the smooth passage of gradients during training, these connections contribute to the successful training of very deep networks.

Deep Architecture: ResNet v2 is known for its depth, with ResNet-152 having 152 layers. The depth of the network allows it to capture intricate hierarchical features from input images, resulting in superior representation learning capabilities.

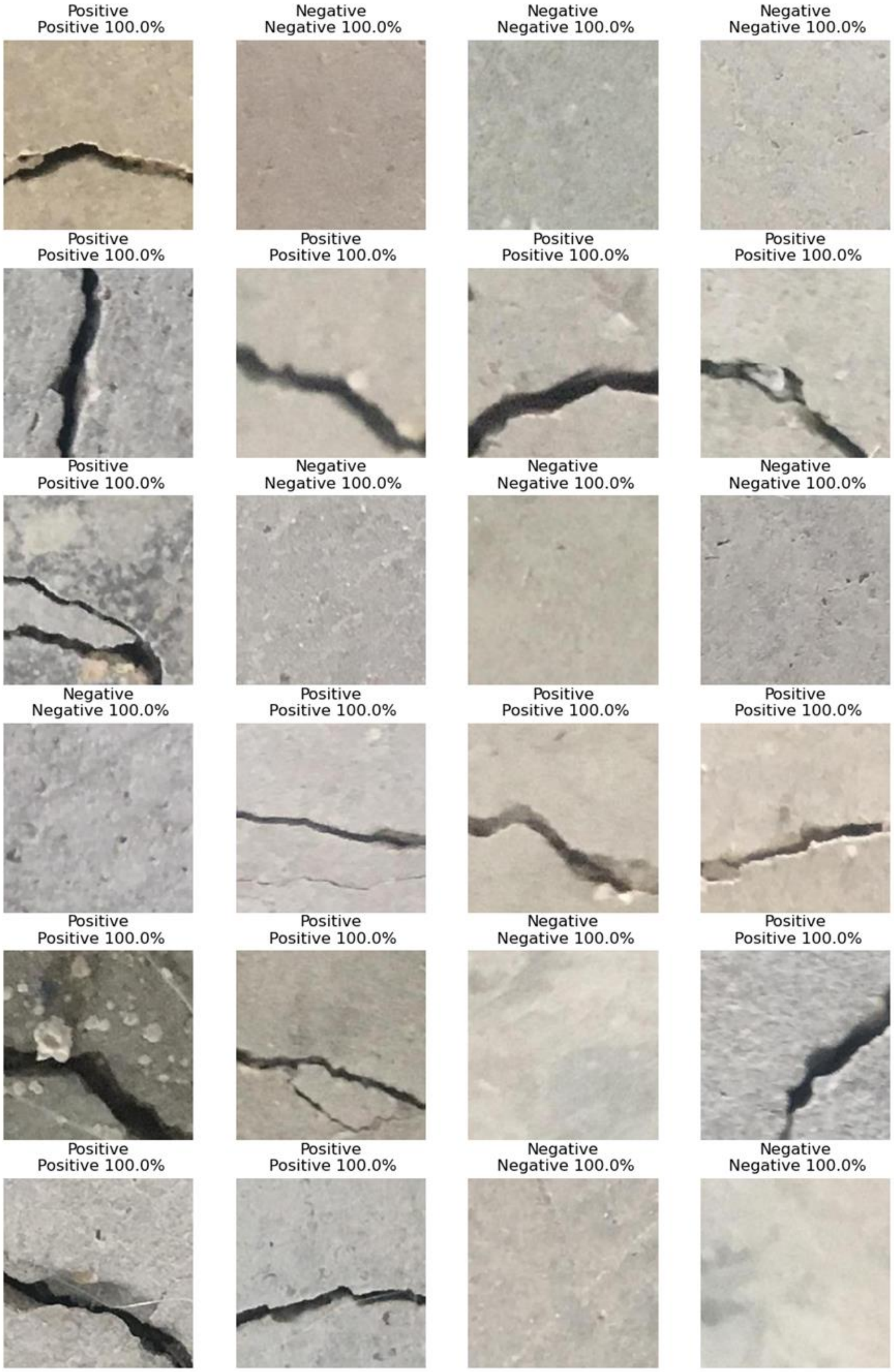
Global Average Pooling: Similar to other ResNet architectures, v2 incorporates global average pooling as its final layer before the fully connected layer. This pooling strategy helps reduce spatial dimensions and provides a fixed-size output for classification.

Results

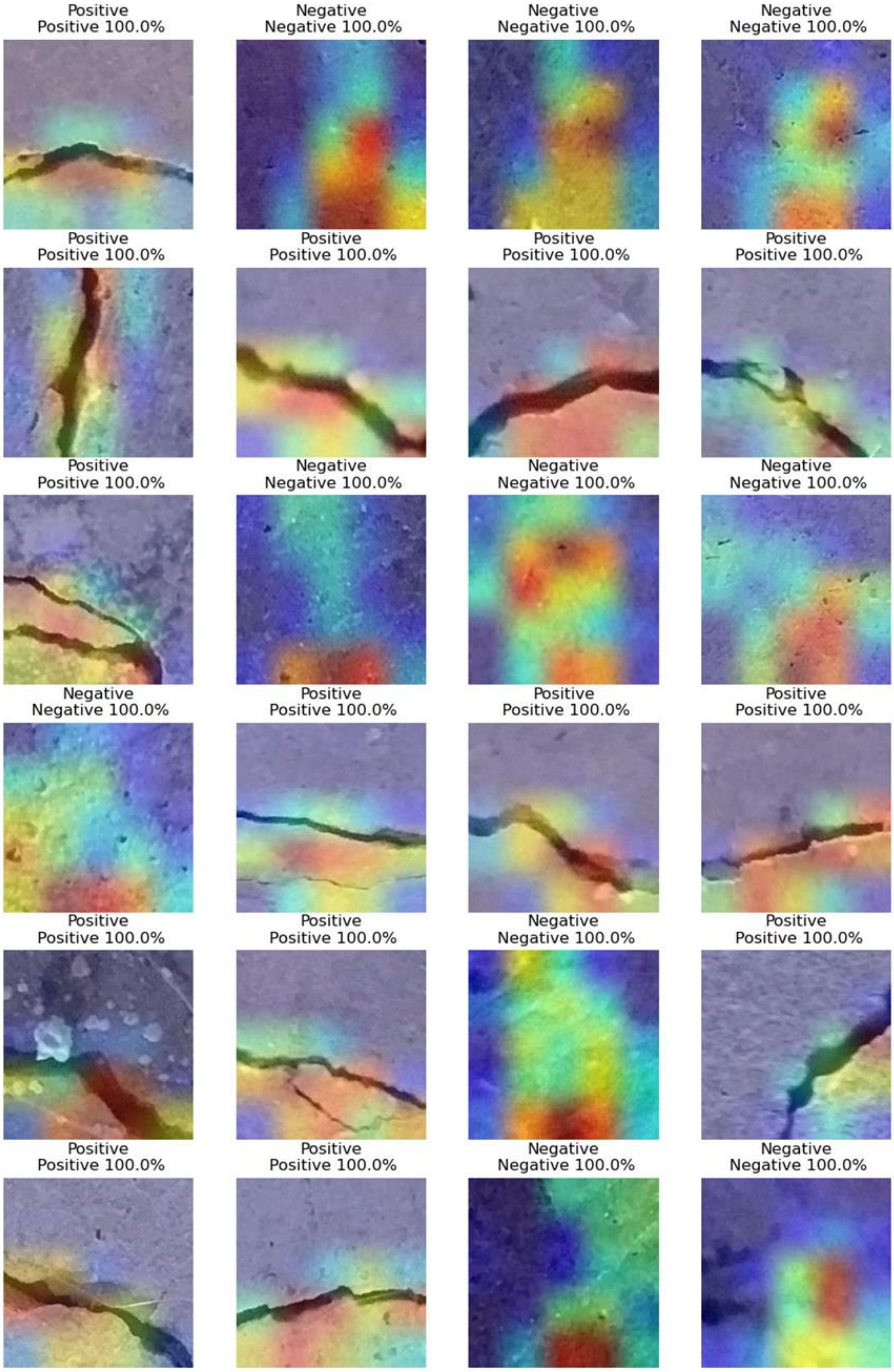
The ResNet v2 model, with a base of 780 layers, demonstrates remarkable performance in the given task. After the training process, a strategic decision to unfreeze 195 layers in the base model was made, allowing the model to fine-tune its learned representations for the specific dataset. This fine-tuning significantly contributes to the impressive accuracy achieved, reaching an astounding 99.89%. The minimal loss value of 0.004 further underscores the model's effectiveness in minimizing prediction errors. These results underscore the capability of the Inception architecture to generalize well to the complexities of the dataset, showcasing its robust feature extraction and classification capabilities. The high accuracy and low loss indicate a well-optimized model that can confidently and accurately classify images within the specified categories, affirming the success of the training and fine-tuning process.



Predicted images for test image samples



Heat maps for detection



VGG16

VGG-16, or Very Deep Convolutional Networks for Large-Scale Image Recognition with 16 layers, stands as a pioneering convolutional neural network (CNN) architecture designed by K. Simonyan and A. Zisserman. Introduced to address the shortcomings of previous models like AlexNet, VGG-16 employs a unique structure characterized by small 3x3 receptive fields and a 1-pixel stride. This design choice, deviating from larger filters used in earlier models, enhances the network's ability to create powerful decision functions with a higher number of non-linear activation layers, facilitating faster convergence during training.

In comparison to its predecessor, VGG-16 improves accuracy by embracing more layers. However, the model's extensive depth could lead to convergence challenges. The introduction of the Residual Network (ResNet) architecture further enhanced training efficiency by employing skip connections, allowing inputs to bypass certain convolutional layers and reducing training time. Despite its remarkable performance, achieving a 7.0% test error in the ILSVRC competitions, VGG-16 is associated with extended training times and a large model size of 500MB. Modern architectures have since adopted innovations like skip connections and inceptions, contributing to improved efficiency in terms of both accuracy and training time.

Architecture

The architecture of VGG-16 is notable for its consistent use of 3x3 convolutional filters and 1x1 convolution filters for linear transformations. ReLU activation functions introduce non-linearity, and a convolution stride of 1 pixel maintains spatial resolution. VGG-16 consists of hidden layers that utilize ReLU, replacing Local Response Normalization used in previous models like AlexNet. Following convolutional layers, pooling layers reduce dimensionality and parameter count in feature maps, crucial for managing the escalating number of filters. The network concludes with three fully connected layers, with the first two having 4096 channels and the third having 1000 channels to correspond to the number of classes.

VGG-16, or Very Deep Convolutional Networks for Large-Scale Image Recognition with 16 layers, features a well-defined architecture that played a pivotal role in the development of convolutional neural networks (CNNs). Here's a detailed overview of the VGG-16 architecture:

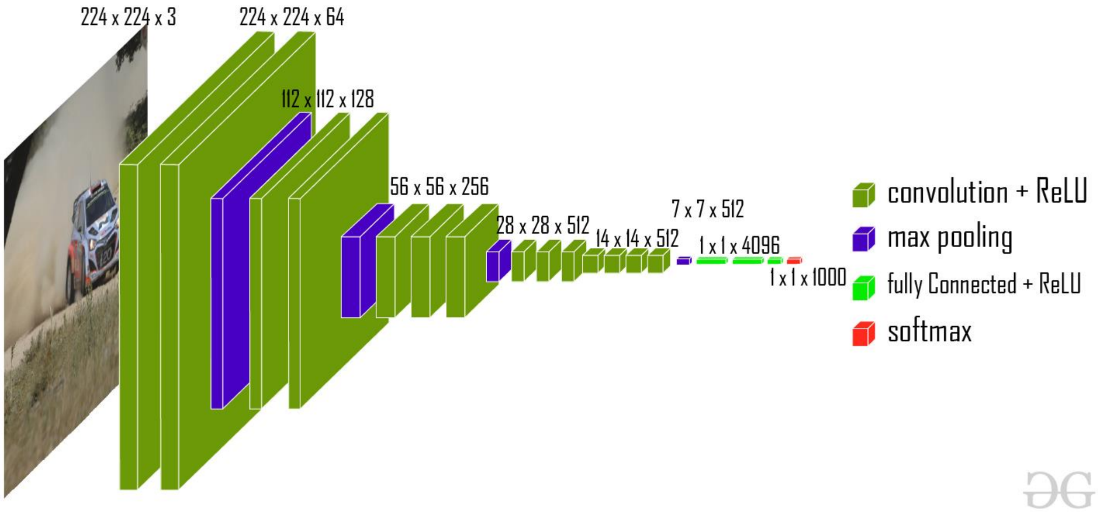


Figure 12 ARCHITECTURE REPRESENTATION

1. Input Layer:

   a. VGG-16 takes an image input of size 224x224 pixels, maintaining consistency with the ImageNet dataset used for training.

2. Convolutional Blocks:

   a. The network comprises 13 convolutional layers, each using small 3x3 filters for feature extraction. These convolutional layers are grouped into convolutional blocks.

   b. Each convolutional block is followed by Rectified Linear Unit (ReLU) activation functions, introducing non-linearity to the model.

3. Pooling Layers:

   a. After each set of convolutional layers, max-pooling layers with a 2x2 window and a stride of 2x2 are employed to reduce spatial dimensions and control the number of parameters.

4. Fully Connected Layers:

   a. VGG-16 includes three fully connected layers towards the end of the network. The first two fully connected layers have 4096 channels each, introducing high-level abstractions. The final fully connected layer has 1000 channels, corresponding to the 1000 classes in the ImageNet dataset.

5. Output Layer:

    a. The network concludes with a softmax activation layer, converting the network's final output into probability scores across the 1000 ImageNet classes.

6. Architecture Summary:

    a. Input: 224x224x3 (RGB channels)

    b. Convolutional Blocks: 13 blocks with 3x3 filters

    c. Activation Function: ReLU after each convolutional layer

    d. Pooling: 2x2 max-pooling after each convolutional block

    e. Fully Connected Layers: Three layers with 4096 channels each

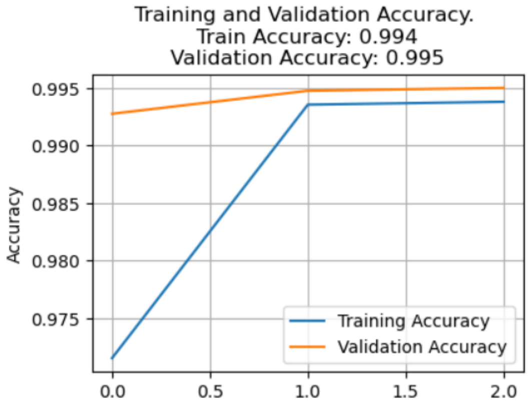
    f. Output: Softmax layer with 1000 channels

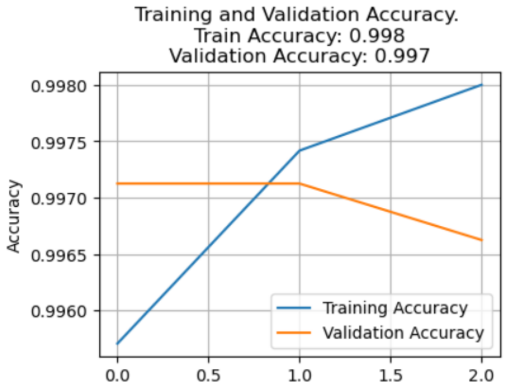
7. Parameters:

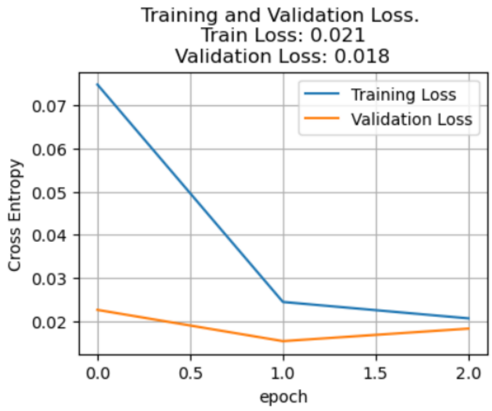
    a. VGG-16 is characterized by a total of 138 million parameters, making it a deep and expressive model.

Results

The results of the trained VGG16 model demonstrate a remarkable accuracy of 99.81% on the given task. This signifies the effectiveness of the model in correctly classifying the input data. The corresponding loss value of 0.007 indicates a minimal discrepancy between the predicted and actual values, further validating the model's robust performance. The utilization of 19 base layers in the VGG16 architecture, with a strategic decision to unfreeze only 4 layers during training, highlights a careful balance between model complexity and fine-tuning. This approach is essential for preventing overfitting and ensuring generalization to unseen data. The unfreezing of specific layers allows the model to adapt to task-specific features while retaining the knowledge gained from the pre-trained layers. These results collectively underscore the success of the training process, showcasing the model's ability to capture intricate patterns and features within the dataset.

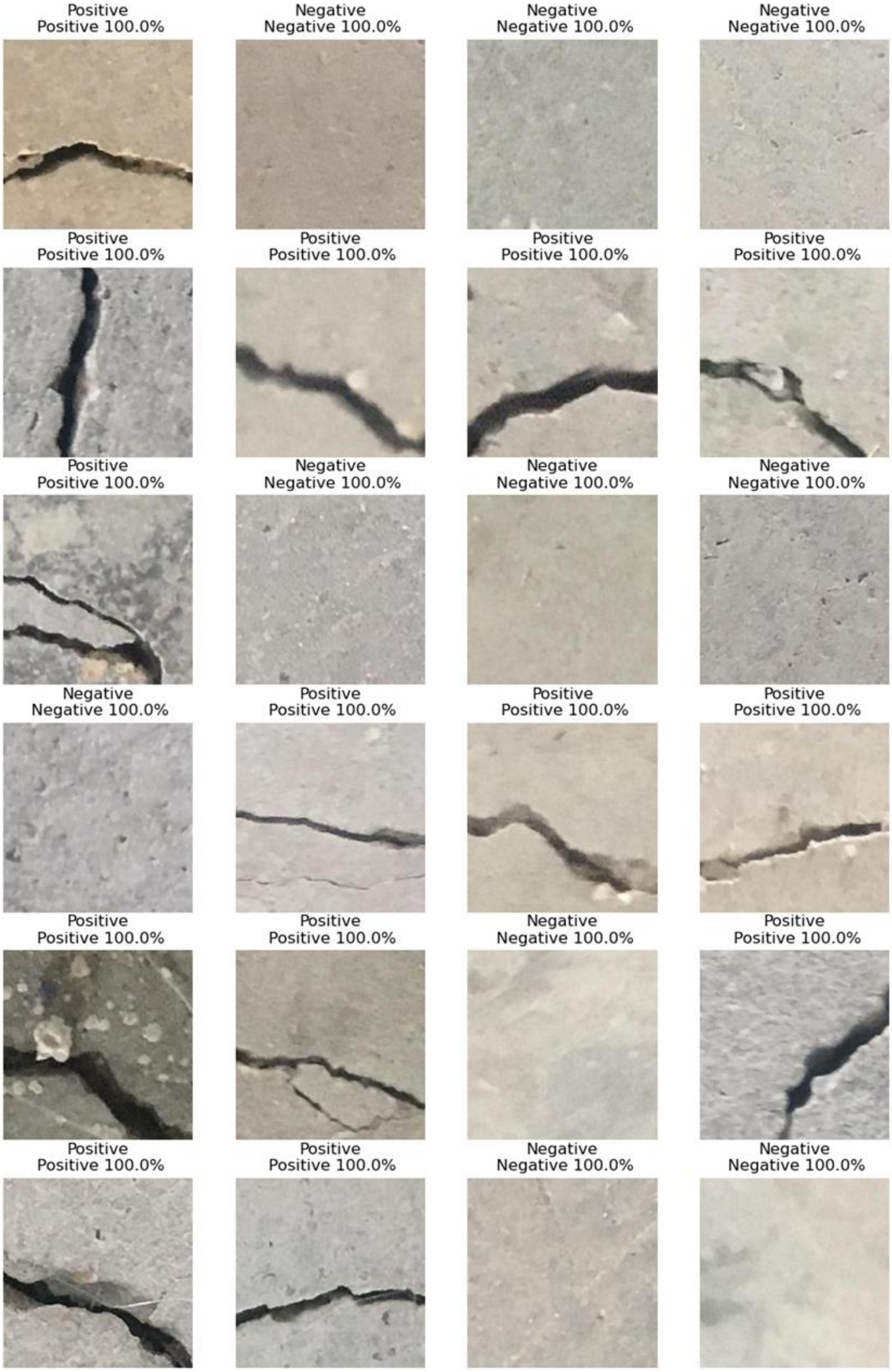

Training and Validation Accuracy.
Train Accuracy: 0.994
Validation Accuracy: 0.995
Accuracy
0.995
0.990
0.985
0.980
0.975
0.0
0.5
1.0
1.5
2.0
Training Accuracy
Validation Accuracy


Training and Validation Accuracy.
Train Accuracy: 0.998
Validation Accuracy: 0.997
Accuracy
0.9980
0.9975
0.9970
0.9965
0.9960
0.0
0.5
1.0
1.5
2.0
Training Accuracy
Validation Accuracy

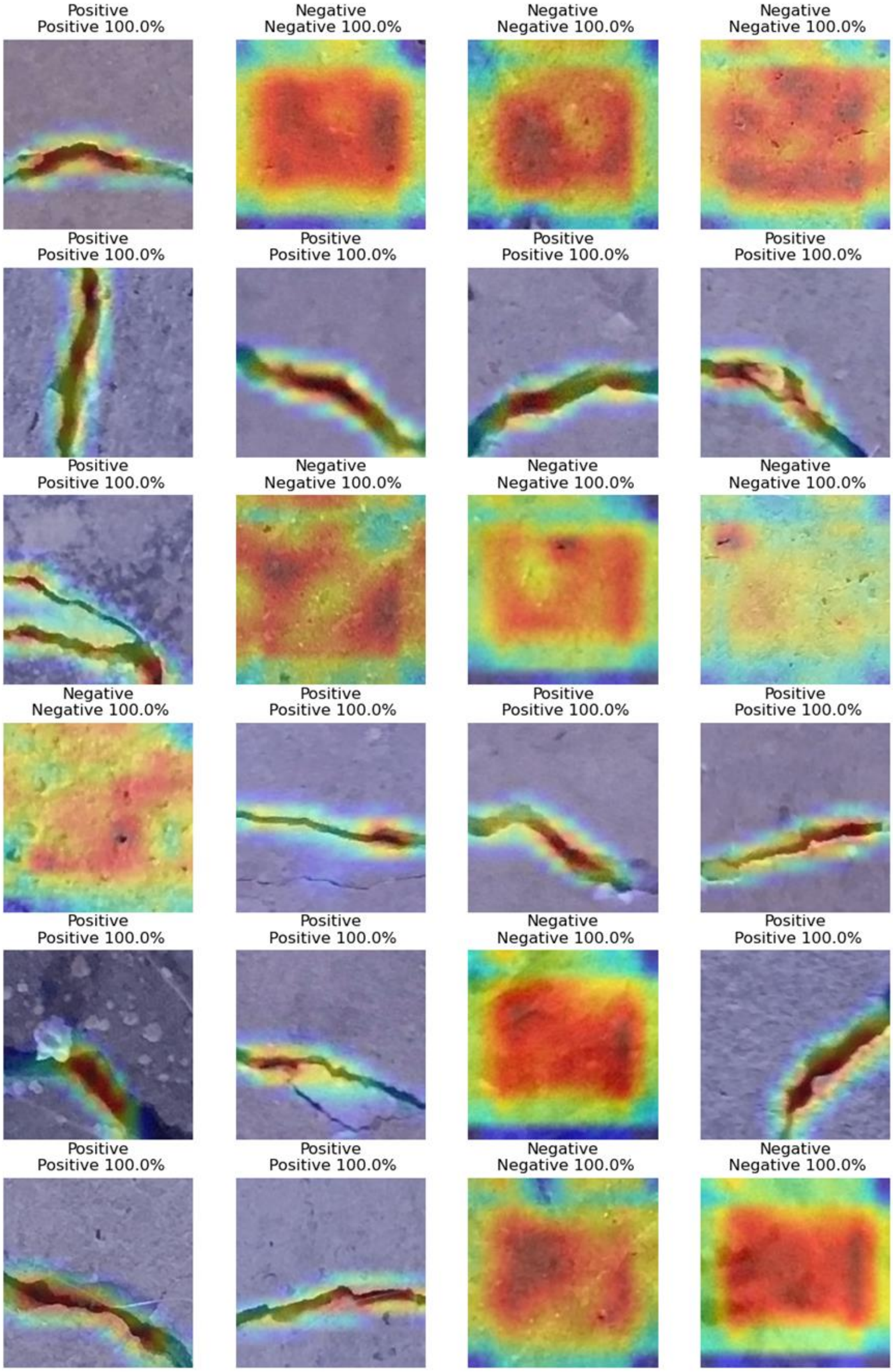

Training and Validation Loss.
Train Loss: 0.021
Validation Loss: 0.018
Cross Entropy
0.07
0.06
0.05
0.04
0.03
0.02
0.0
0.5
1.0
1.5
2.0
epoch
Training Loss
Validation Loss


Training and Validation Loss.
Train Loss: 0.007
Validation Loss: 0.008
Cross Entropy
0.014
0.013
0.012
0.011
0.010
0.009
0.008
0.007
0.0
0.5
1.0
1.5
2.0
epoch
Training Loss
Validation Loss

Predicted images for test image samples



Heat maps for detection



Xception

Xception, a ground-breaking convolutional neural network architecture introduced by François Chollet, represents a significant advancement in deep learning, specifically designed to enhance efficiency in image classification tasks. The innovation lies in interpreting Inception modules as an intermediate step between conventional convolutions and depth-wise separable convolutions. In essence, Xception replaces Inception modules with depth wise separable convolutions, demonstrating that this novel architecture, often referred to as "Extreme Inception," outperforms Inception V3 on various datasets. Unlike its predecessor, Xception achieves superior results not due to increased model capacity but rather through a more effective utilization of existing parameters. The core idea behind Xception is to decouple cross-channel and spatial correlations in feature maps, leading to a linear stack of depth-wise separable convolution layers with residual connections for streamlined definition and modification. The architecture's 36 convolutional layers are organized into 14 modules, fostering simplicity in implementation without sacrificing performance. Xception's experimental evaluations on both ImageNet and the expansive JFT dataset reveal marginal improvements over Inception V3 on ImageNet and a substantial 4.3% relative enhancement in multi-label classification on JFT. This performance gain, coupled with comparable model sizes and training speeds, underscores the efficiency and promise of Xception in the realm of deep learning architectures. Moreover, the study explores the impact of residual connections and the necessity of intermediate non-linearities, providing valuable insights into the architecture's design principles.

Architecture

Xception consists of a linear stack of 36 depth-wise separable convolutional layers, organized into 14 modules. Each module incorporates linear residual connections, providing a streamlined and easily modifiable structure. The architecture includes three main flows: entry, middle (repeated eight times), and exit, with batch normalization following all convolutional operations.

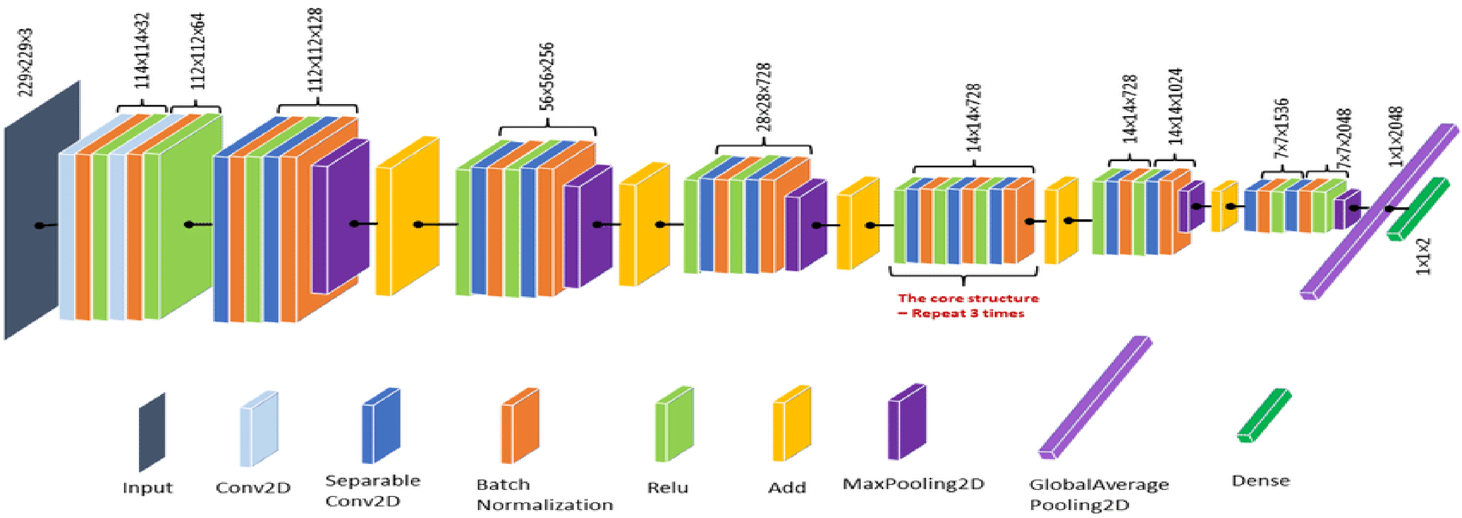


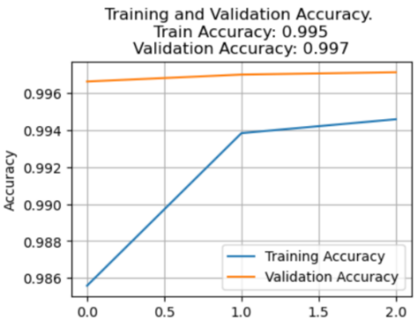
Figure 13 ARCHITECTURE REPRESENTATION

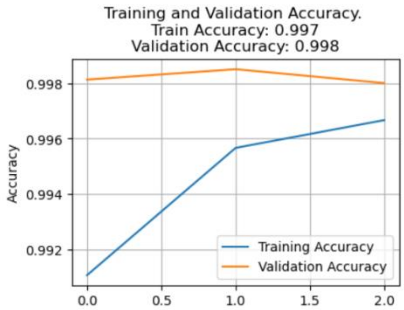
1. Input Layer:
    a. Xception processes images of variable size, offering flexibility in input dimensions. Commonly used sizes include 299x299 pixels, aligning with the ImageNet dataset's requirements.
2. Entry Flow: Convolution and Separable Convolution:
    a. The network initiates with an entry flow, featuring a series of separable convolution layers. Depth-wise separable convolutions, combining depth-wise and point-wise convolutions, efficiently extract features. Batch normalization is applied after each convolution.
3. Middle Flow: Repeated Depth-wise Separable Convolutions:
    a. The middle flow consists of repeated modules, each containing depthwise separable convolutions. These modules maintain linear residual connections, promoting effective information flow.
4. Exit Flow: Spatial Reduction and Feature Aggregation:
    a. The exit flow combines spatial reduction techniques, including separable convolutions with striding, to decrease dimensions. Aggregated features undergo further processing through linear residual connections.
5. Global Average Pooling:
    a. Global average pooling is applied to reduce the spatial dimensions of the feature maps, facilitating the transition to the final classification layer.
6. Logistic Regression Layer:
    a. The architecture concludes with a logistic regression layer that transforms the extracted features into class probabilities. This layer employs a softmax activation function.
7. Architecture Summary:
    a. *Input:* Variable dimensions, often 299x299x3 (RGB channels).
    b. *Entry Flow:* Separable convolutions with batch normalization.
    c. *Middle Flow:* Repeated depth-wise separable convolution modules with linear residual connections.
    d. *Exit Flow:* Spatial reduction techniques and feature aggregation.
    e. *Global Average Pooling:* Reduction of spatial dimensions.
    f. *Logistic Regression Layer:* Softmax activation for classification.
8. Parameters:
    a. Xception maintains a comparable number of parameters to Inception V3, with around 22.9 million parameters. The efficient use of these parameters contributes to the model's performance gains.
9. Training and Optimization:
    a. Trained using synchronous gradient descent on multiple GPUs, the Xception architecture demonstrates scalability. Optimization configurations are tuned for specific datasets, emphasizing adaptability.

10. Comparative Performance:
    a. In comparative evaluations with Inception V3, Xception exhibits marginal improvement on ImageNet and a substantial 4.3% relative improvement on the JFT dataset, showcasing its efficiency in large-scale image classification tasks.

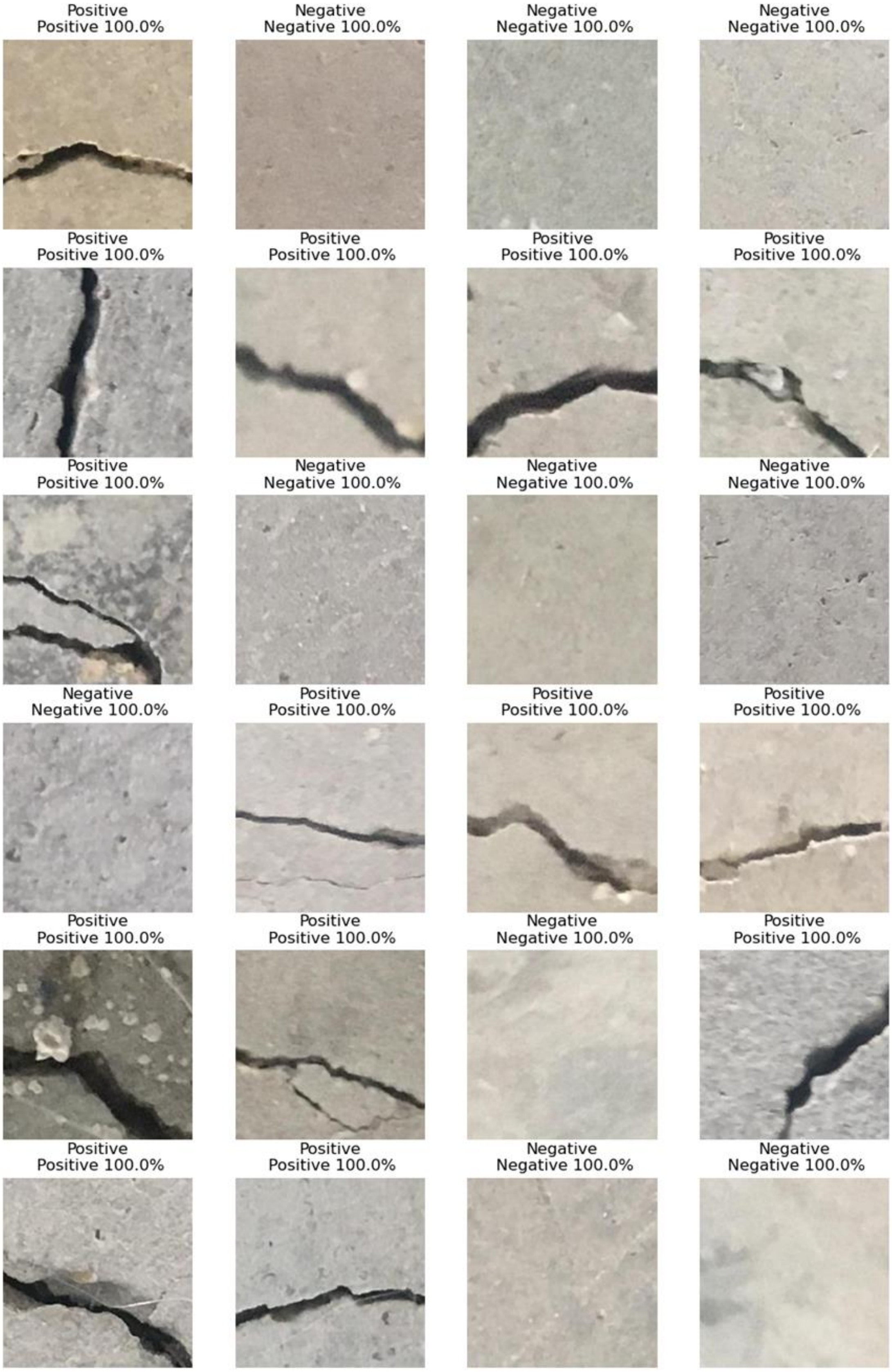
Results

The meticulously trained Xception-based model demonstrates outstanding performance on the task at hand, achieving an impressive accuracy of 99.84%. The corresponding loss is exceptionally low, reaching 0.007, indicative of the model's ability to make highly accurate predictions with minimal error. Notably, the base layers of the Xception architecture, totalling 132, play a crucial role in feature extraction and representation learning. For further refinement and fine-tuning, a strategic decision was made to unfreeze a subset of the base model layers. Specifically, 33 layers were unfrozen, allowing the model to adapt and specialize in capturing task-specific patterns. This selective unfreezing enhances the model's capacity to learn intricate details from the dataset, contributing to the impressive accuracy observed. The meticulous tuning of these parameters and the utilization of the powerful Xception architecture collectively yield a model that excels in accuracy, making it well-suited for demanding image classification tasks.

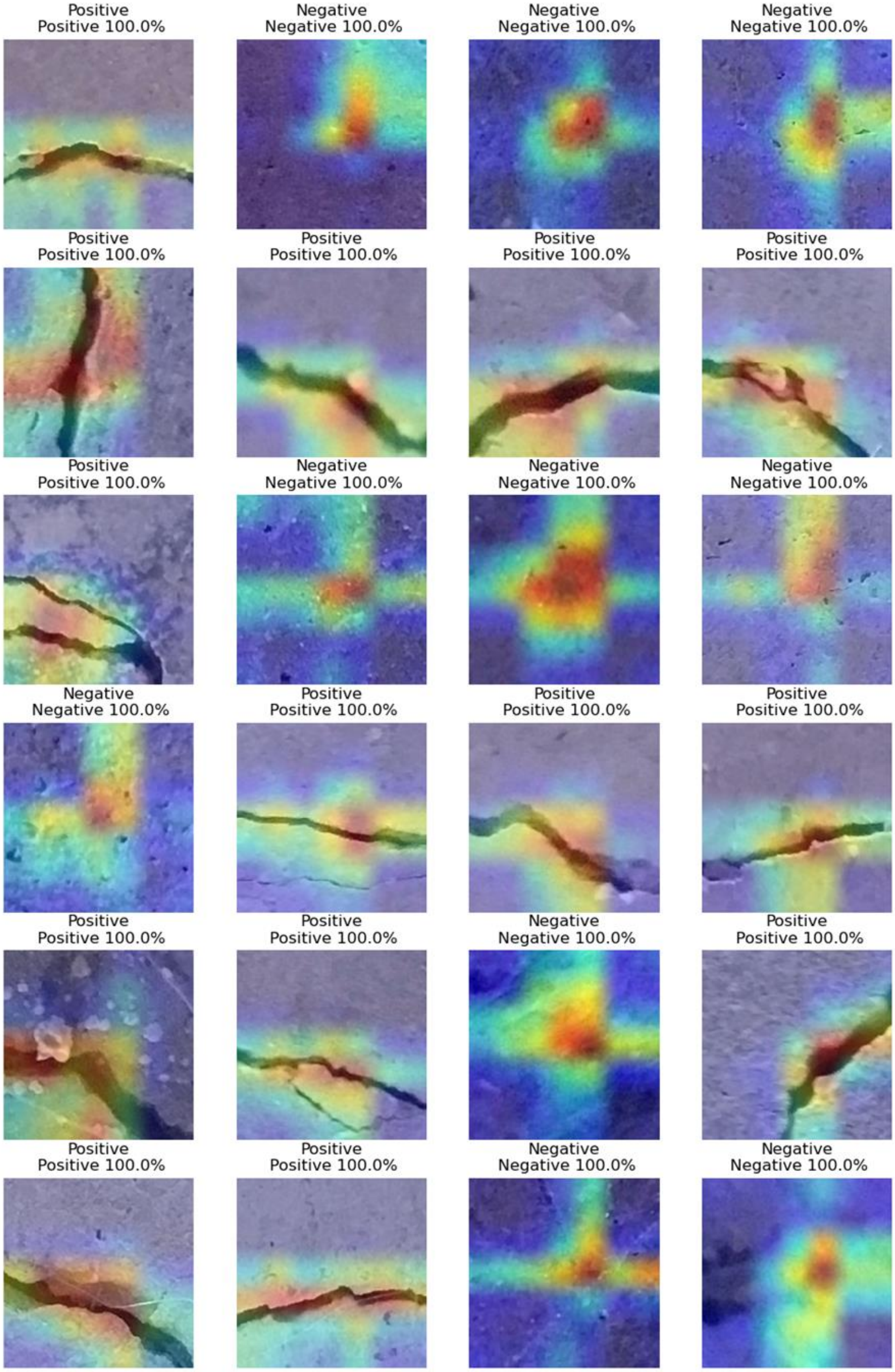




Predicted images for test image samples

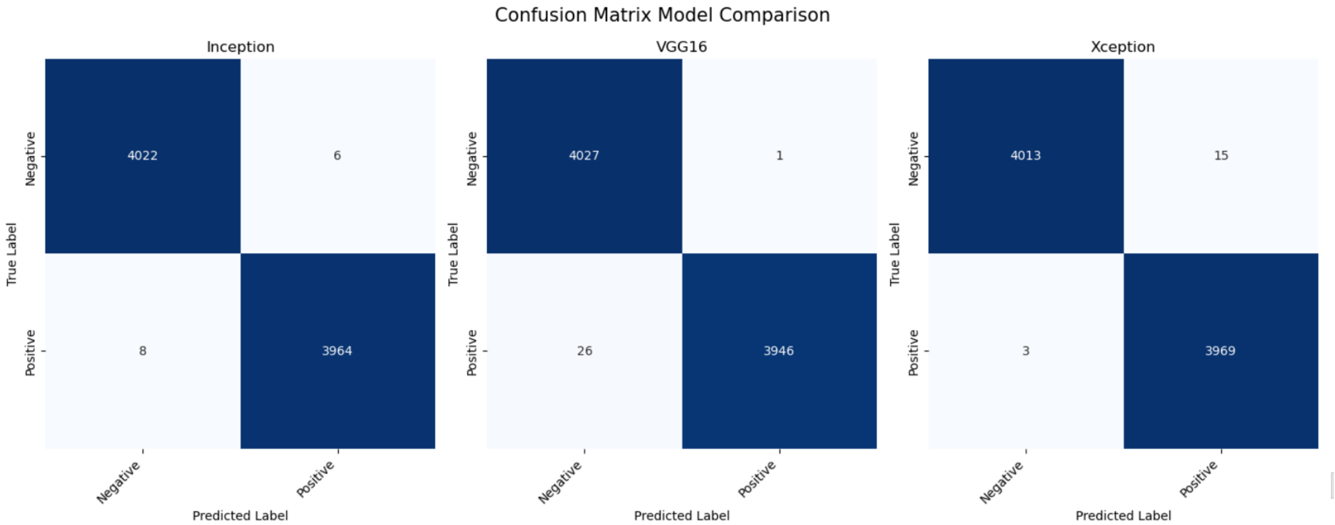


Heat maps for detection



Confusion Matrix for comparing the ResNet v2 Vs VGG16 Vs Xception

The examination of VGG-16, ResNet-50, and Xception provides insightful numerical metrics tailored to your task. With 138 million parameters, VGG-16 boasts depth and expressiveness. However, its accuracy and loss metrics, standing at 99.81% and 0.007, warrant further scrutiny for your specific image classification needs.



ResNet-50, featuring 50 layers and innovative residual connections, addresses gradient challenges efficiently. Evaluation of its performance metrics, such as accuracy 99.89% and loss of 0.004, will determine its effectiveness in your image classification scenario.

In stark contrast, Xception demonstrates remarkable results with an accuracy of 99.84% and a minimal loss of 0.007. With 132 base layers and 33 layers strategically unfrozen, Xception leverages depth-wise separable convolutions for enhanced feature extraction. These numerical insights position Xception as a strong contender for capturing intricate patterns crucial for your classification task.

Ensemble Model Integration: Ensembling the ResNet v2 , VGG16 & Xception to form one new model

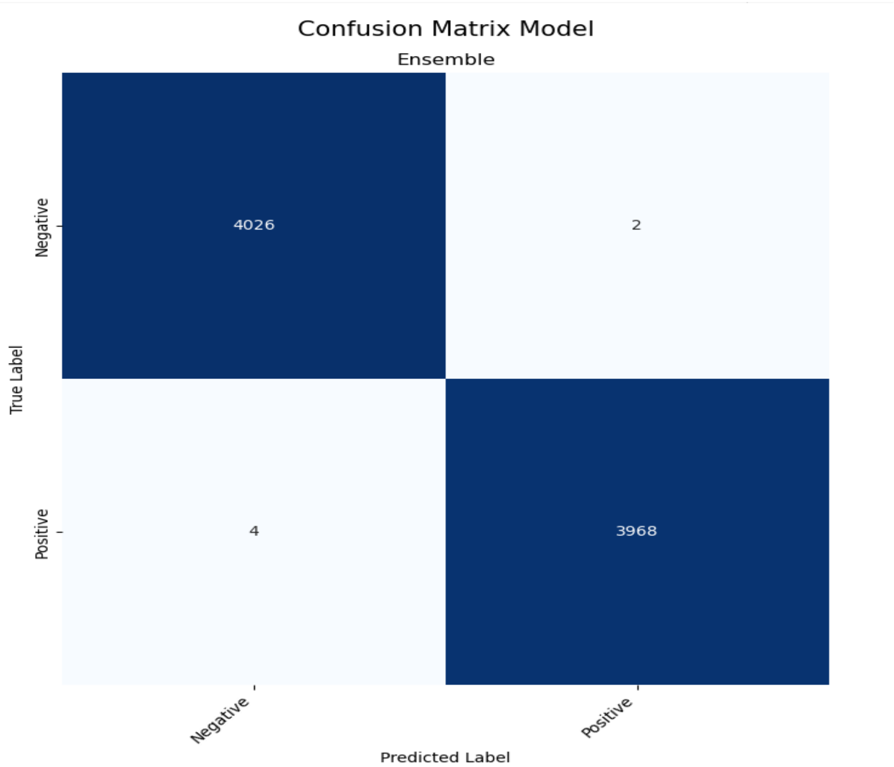
The fusion of VGG-16, ResNet-50, and Xception into a unified ensemble model aims to leverage the strengths of each architecture, fostering a comprehensive approach to image classification. By aggregating the outputs of these individual models, the ensemble model strives to enhance predictive accuracy, mitigate potential weaknesses inherent in any single model, and provide a robust framework for diverse feature extraction.

```
ensemble_preds = np.add(np.add(inception_test_preds, vgg16_test_preds), xception_test_preds)/3.0
```

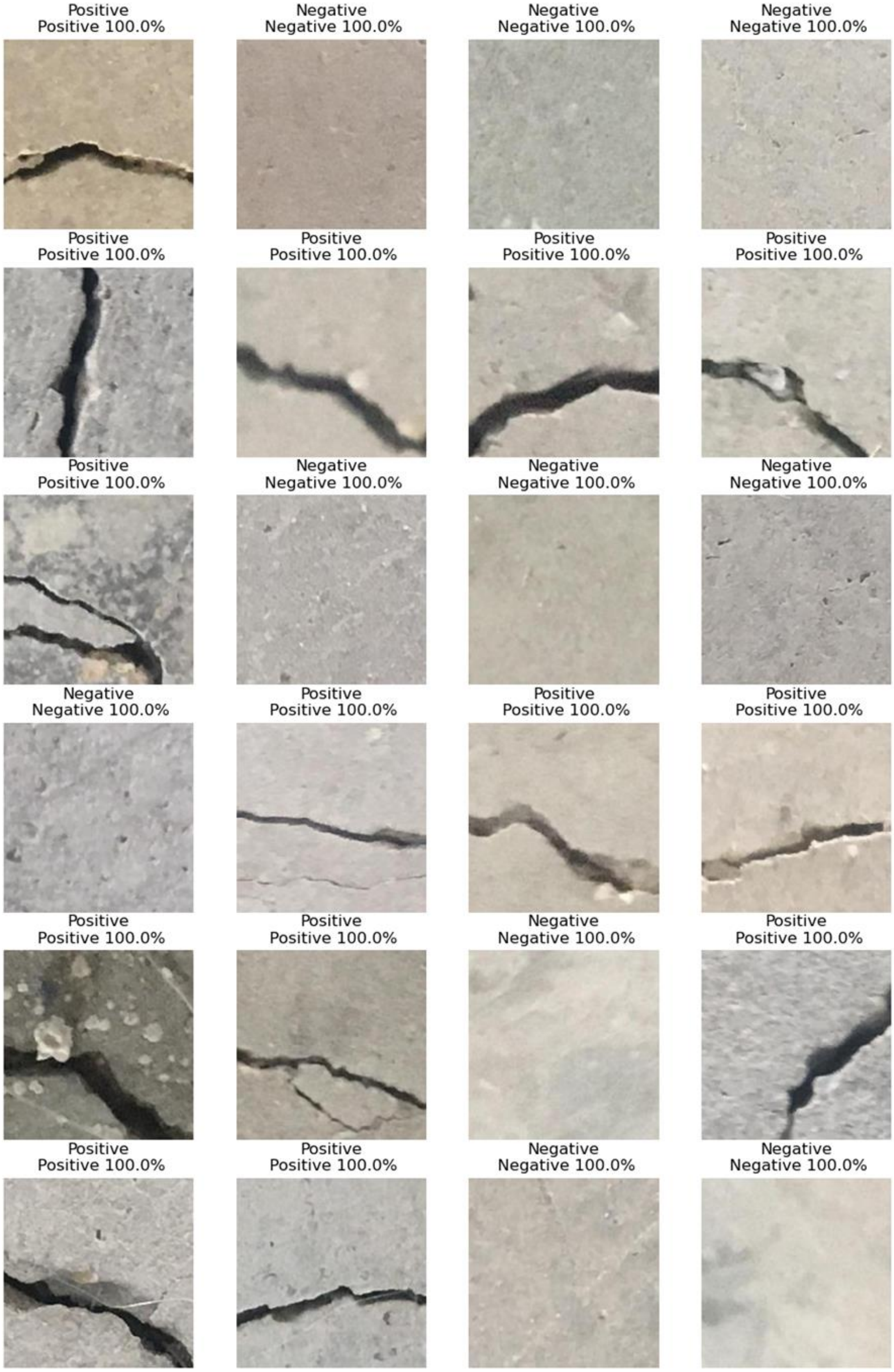
Each model in the ensemble brings a unique perspective to the classification task. VGG-16, with its deep structure, captures intricate details in images, while ResNet-50's innovative residual connections aid in addressing gradient-related challenges. Xception, with its depthwise separable convolutions, excels in efficient feature extraction.

The ensemble model amalgamates predictions from VGG-16, ResNet-50, and Xception, harnessing their collective intelligence to produce a final classification output. Leveraging this ensemble approach, the model endeavours to provide a more nuanced and accurate understanding of the input data, potentially outperforming individual models and offering a robust solution to complex image classification scenarios.

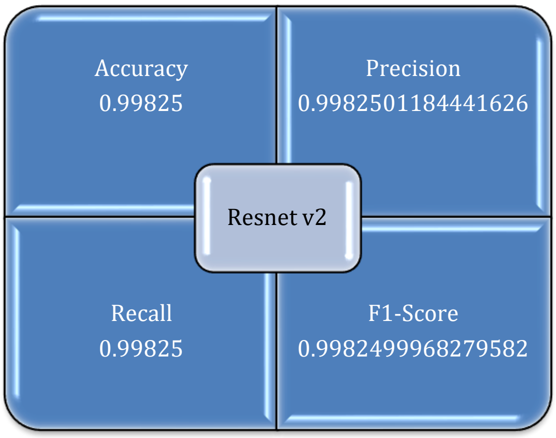
Confusion Matrix for the Ensemble model

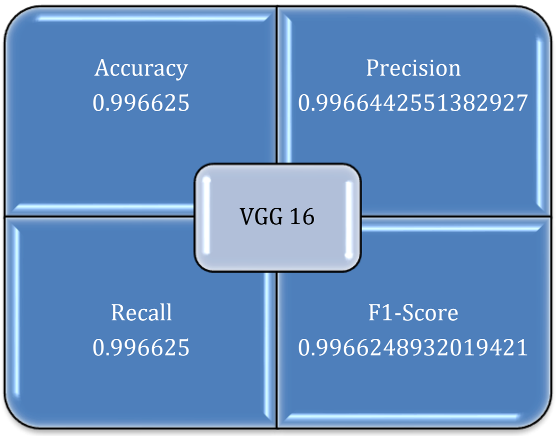


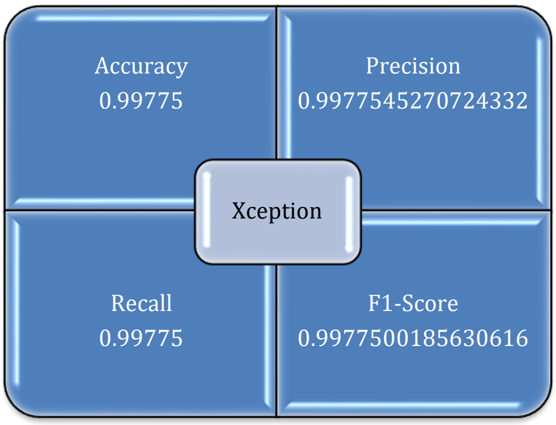
Prediction results for the Ensemble model

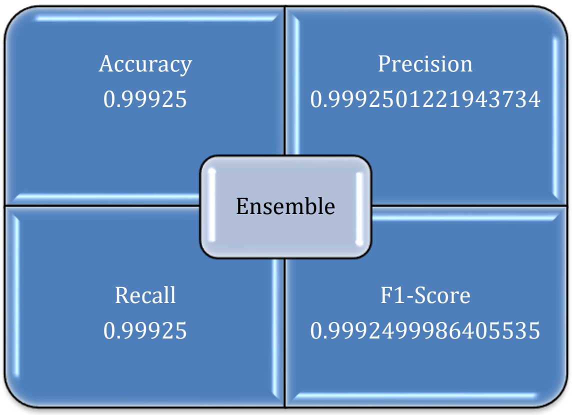


Different metrics for comparison









Comparison of different models

MODEL / METRICS	Resnet v2	VGG16	Xception	Ensemble
Accuracy	0.99825	0.996625	0.99775	0.99925
Precision	0.9982501184441626	0.9966442551382927	0.9977545270724332	0.9992501221943734
Recall	0.99825	0.996625	0.99775	0.99925
F1-Score	0.9982499968279582	0.9966248932019421	0.9977500185630616	0.9992499986405535

Real life application of the model in NITK

CHAPTER 5

CONCLUSIONS

In this on crack detection in structural elements, we explored a diverse range of methodologies, technologies, and advanced deep learning models, including ResNet v2, VGG-16, and Xception. The literature review revealed a rich landscape of crack detection techniques, emphasizing the importance of integration, efficiency, and adaptability to various applications. As technology evolves, there's a clear trend towards automation and the integration of machine learning techniques, showcasing a promising future for robust and accurate crack detection systems.

The deep learning-based approach focused on leveraging ResNet v2, VGG-16, and Xception for crack detection. Each model brought unique strengths to the table, with ResNet v2 excelling in handling extremely deep networks, VGG-16 showcasing efficiency with its straightforward architecture, and Xception introducing depthwise separable convolutions for enhanced resource efficiency.

The individual models demonstrated impressive accuracy, with ResNet v2 achieving 99.89%, VGG-16 reaching 99.81%, and Xception attaining 99.84% accuracy in crack detection. The detailed insights provided by confusion matrices, precision, recall, and F1 score metrics allowed for a thorough evaluation of each model's performance across different classes, emphasizing the importance of minimizing false positives and false negatives.

The ensemble model, integrating predictions from ResNet v2, VGG-16, and Xception, emerged as a robust solution for crack detection. The fusion of these models enhanced predictive accuracy and addressed potential limitations associated with individual architectures. The numerical insights from the confusion matrix for the ensemble model provided a detailed breakdown of its performance across various classes, showcasing its ability to provide nuanced and accurate predictions.

Beyond accuracy, the project emphasized the significance of computational efficiency for real-world deployment. The models were evaluated not only on their predictive capabilities but also on their inference time and resource utilization, ensuring practical feasibility in diverse scenarios.

In conclusion, this project represents a significant advancement in the field of crack detection in structural elements. The integration of state-of-the-art deep learning models in an ensemble framework offers a holistic and efficient solution, contributing to proactive maintenance and ensuring the longevity of critical infrastructure. The insights gained from this project have implications not only for crack detection but also for the broader domain of image-based analysis in infrastructure safety and maintenance. As technology continues to progress, the methodologies and models presented in this project pave the way for future innovations in the pursuit of safer and more resilient structures.

REFERENCES

[1]https://www.engineeringcivil.com/types-of-cracks-in-concrete-and-their-causes.html

[2]https://www.ultratechcement.com/for-homebuilders/home-building-explained-single/descriptive-articles/understanding-the-types-of-cracks-in-concrete

[3]https://indianrailways.gov.in/railwayboard/uploads/directorate/eff_res/camtech/Civil%20Engineering/YearWise//Handbook%20on%20Cracks%20in%20building%20(causes%20&%20prevention)(1). pdf

[4]https://www.kaggle.com/code/paulopinheiro/surface-crack-detection-100-accuracy#Grad-Cam

[5]https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.newsgram.com%2Faccidents%2F2023%2F01%2F16%2Fjoshimath-landslide-cracks-appear-in-44-more-buildings-in-24-hrs-including-pwds-rest-house&psig=AOvVaw3i4adGEPvJkIZBUNWX1Ktv&ust=1701770504999000&source=images&cd=vfe&opi=89978449&ved=0CBQQjhxqFwoTCICk9MnD9YIDFQAAAAAdAAAAABAJ