REPORT OF PRE-TRAINED MODEL

Tensorflow

SUMMARY

The report highlights the comparative analysis of machine learning models on image recognition tasks, emphasizing the critical balance between accuracy, computational efficiency, and model size necessary for optimal application performance.

Ali Huseynov

AI engineer…

https://betterprogramming.pub/image-classification-with-pre-trained-models-baecb2019973

The examination of six cutting-edge machine learning models—ResNet152V2, ResNet50, DenseNet, MobileNet, VGG16, and NASNet—reveals a nuanced landscape of options available for image classification tasks, each balancing computational efficiency, model size, and accuracy in unique ways.

Three Python files were uploaded to GitHub, each designed to fulfill distinct tasks within a machine learning or data analysis workflow:

1. **5Model.py**: This script is tailored for comparing five different models, showcasing their capabilities by providing the top five predictions from each. It's a comparative tool that allows for an in-depth evaluation of how each model performs on the same dataset or input, highlighting the prediction diversity among different models.
2. **Combined.py**: Engineered to operate with the same five models, this script not only compares the results but also delves into the accuracy percentages and detailed predictions of each model. It offers a nuanced view by revealing which models perform better in terms of prediction accuracy and provides full insights into the predictions made by each model.
3. **ForAPI.py**: Shares a core codebase with Combined.py, but is streamlined for a different output format. While Combined.py presents model names alongside their predictions, ForAPI.py is designed to output only the predictions and their confidence percentages, omitting the model names. This simplification is particularly suited for API responses where the emphasis is on the prediction details rather than the specifics of the model making those predictions.

Model	Size	Top-1 Accuracy	Top-5 Accuracy	Parameters	Depth
Xception	88 MB	0.790	0.945	22,910,480	126
VGG16	528 MB	0.713	0.901	138,357,544	23
VGG19	549 MB	0.713	0.900	143,667,240	26
ResNet50	98 MB	0.749	0.921	25,636,712	–
ResNet101	171 MB	0.764	0.928	44,707,176	–
ResNet152	232 MB	0.766	0.931	60,419,944	–
ResNet50V2	98 MB	0.760	0.930	25,613,800	–
ResNet101V2	171 MB	0.772	0.938	44,675,560	–
ResNet152V2	232 MB	0.780	0.942	60,380,648	–
ResNeXt50	96 MB	0.777	0.938	25,097,128	–
ResNeXt101	170 MB	0.787	0.943	44,315,560	–
InceptionV3	92 MB	0.779	0.937	23,851,784	159
InceptionResNetV2	215 MB	0.803	0.953	55,873,736	572
MobileNet	16 MB	0.704	0.895	4,253,864	88
MobileNetV2	14 MB	0.713	0.901	3,538,984	88
DenseNet121	33 MB	0.750	0.923	8,062,504	121
DenseNet169	57 MB	0.762	0.932	14,307,880	169
DenseNet201	80 MB	0.773	0.936	20,242,984	201
NASNetMobile	23 MB	0.744	0.919	5,326,716	–
NASNetLarge	343 MB	0.825	0.960	88,949,818	–

https://image-net.org/challenges/LSVRC/2014/browse-synsets You can find classes in this link.

https://m.media-amazon.com/images/I/71KwPy8BPiL._AC_UF1000,1000_QL80_.jpg

https://media.istockphoto.com/id/155439315/photo/passenger-airplane-flying-above-clouds-during-sunset.jpg?s=612x612&w=0&k=20&c=LJWadbs3B-jSGJBVy9s0f8gZMHi2NvWFXa3VJ2lFcL0=

https://res.cloudinary.com/padi/image/upload/v1701290031/CDN/commerce/wreck-diver.jpg

https://c1.wallpaperflare.com/preview/826/375/678/race-car-sport-car-sport-speed.jpg

https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcQEysfzzeDbVyV9s9Atx16tFld4snlaCbvGFg&usqp=CAU

https://avatars.mds.yandex.net/i?id=7be615771fa05f41231c53d56cefa446eda3f067-10178110-images-thumbs&n=13

https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcQEysfzzeDbVyV9s9Atx16tFld4snlaCbvGFg&usqp=CAU

https://idfg.idaho.gov/sites/default/files/styles/article_header/public/field/image/34a99d1c-75c1-4035-8dbe-76546f83b279.jpg?itok=OCpTFYy6

https://www.freecodecamp.org/news/content/images/2021/11/niclas-illg-wzVQp_NRIHg-unsplash.jpg}

https://images2.minutemediacdn.com/image/upload/c_fill,w_1440,ar_16:9,f_auto,q_auto,g_auto/shape/cover/sport/mirror-gettyimages-537282880-0-c1e4c2d1b9d7ff400d4cc8fce2301f43.jpg

https://smarthistory.org/wp-content/uploads/2019/07/cordoba.jpg

These are 10 links of pictures that I used my codes, and these are results:
Introduction: This report provides a comprehensive analysis of five different image classification models: ResNet, VGG, MobileNet, NASNet, and DenseNet. By examining the models' predictions across various images, we will assess their accuracy, strengths, and weaknesses. Additionally, the report will include insights into the relative sizes and computational efficiencies of these models, important factors in model selection for specific applications.

Model Overview:

1. **ResNet (Residual Network):** Known for introducing residual learning blocks to alleviate the vanishing gradient problem, enabling the training of much deeper networks. It is widely used due to its ability to achieve high accuracy with relatively efficient computation.
2. **VGG (Visual Geometry Group):** Characterized by its simplicity, using only 3x3 convolutional layers stacked on top of each other in increasing depth. It is heavier and slower than more modern architectures but remains popular due to its high accuracy in classification tasks.
3. **MobileNet:** Designed for mobile and embedded vision applications. MobileNet utilizes depthwise separable convolutions to reduce model size and complexity, optimizing for speed and efficiency with minimal loss in accuracy.
4. **NASNet (Neural Architecture Search Network):** Uses reinforcement learning to discover the best architecture for the task, leading to highly efficient models in terms of both size and computational resource requirements.
5. **DenseNet (Densely Connected Convolutional Networks):** Features densely connected layers where each layer is connected to every other layer in a feed-forward fashion. This architecture ensures maximum information flow between layers, making it highly efficient in parameter use.

Analysis of Model Predictions Across Different Images:

- **Soccer Ball Image Analysis:**
  - MobileNet stands out with an exceptional 99.12% accuracy in identifying the soccer ball, showcasing its strength in mobile and embedded applications where accuracy and efficiency are paramount.
  - ResNet and NASNet also showed high accuracy, indicating their robustness in handling complex patterns.
- **Airplane Image Analysis:**
  - The diversity in predictions, with some models confusing the aircraft for a "space shuttle," highlights the challenge of classifying objects with similar features. MobileNet and VGG favored "airliner," suggesting a better handling of context cues.
- **Underwater Wreck Image Analysis:**
  - MobileNet's 92.96% accuracy in identifying the wreck demonstrates its capability in distinguishing details in complex environments, potentially beneficial for applications in underwater exploration.
- **Race Car Image Analysis:**
  - The high accuracy of MobileNet and DenseNet in identifying the race car aligns with their architectural strengths, where detail-oriented and high-parameter efficiency is crucial.
- **Furniture Image Analysis:**
  - The diversity in object identification across the models, from "entertainment_center" to "bookcase," reflects the nuanced understanding required for indoor scenes, with MobileNet showing high accuracy for bookcase identification.
- **Animal Image Analysis:**

    - NASNet's superior identification of a "lion" in one of the images underscores the model's adaptability and architectural efficiency, likely benefiting from its neural architecture search foundation.
- **Fish Image Analysis:**
    - The consensus on "sturgeon" across models, especially the near-perfect accuracy by MobileNet, highlights the effectiveness of current model architectures in classifying distinct biological features.
- **Computer Setup Image Analysis:**
    - The consistent high accuracy in identifying "desktop_computer" illustrates the models' capabilities in recognizing technology-related objects, with NASNet showing noteworthy efficiency.
- **Car Mirror Image Analysis:**
    - The unanimous high accuracy in identifying "car_mirror," especially by MobileNet and DenseNet, reflects their strength in dealing with highly reflective and shape-specific objects.
- **Architectural Structure Image Analysis:**
    - Variations in predictions, such as "vault," "throne," and "altar," highlight the challenges in architectural and historical classification, with NASNet and DenseNet providing intriguing insights into their architectural learning capabilities.

Conclusion: This comparison underscores the advancements in image classification models, revealing their specialized capabilities and potential applications. MobileNet emerges as notably efficient for mobile applications, balancing accuracy with computational resource constraints. DenseNet and NASNet offer promising avenues for high accuracy and efficiency, particularly in distinguishing complex patterns and textures. Meanwhile, ResNet and VGG, despite their older architectures, continue to provide robust frameworks for a broad spectrum of image recognition tasks.

The choice of model ultimately depends on the specific requirements of the task, including the balance between accuracy, speed, and computational resources. Future research and development will likely focus on enhancing models' efficiency and adaptability, potentially incorporating advances in neural architecture search and deep learning optimizations.

Addendum to Image Classification Model Analysis Report: Ali's Recommendation for ResNet50

User Recommendation: Based on the analysis and the specific outcomes derived from the assortment of images tested, the user advises the adoption of the ResNet50 model for image classification tasks. This recommendation is predicated on ResNet50's particularly detailed and accurate predictions across a wide variety of images.

Rationale:

- **Depth and Complexity Handling:** ResNet50's architecture, known for its "deep" structure facilitated by residual connections, allows for the training of models that are both deep and capable of handling the vanishing gradient problem. This enables it to learn from a vast variety of features and details that might be missed by shallower networks.
- **Balanced Performance:** While not always the leader in accuracy in every single case, ResNet50 consistently provides high-accuracy predictions across a diverse set of

images. This balance makes it a reliable choice for applications where a wide variety of objects needs to be accurately recognized.

- **Specialized Detail Recognition:** The user points out that ResNet50 often gives "specially detailed outputs," underscoring its ability to not just identify objects in images but to do so with a high level of precision and attention to detail. This capability is particularly valuable in scenarios where fine distinctions between similar objects are crucial.
- **Versatility:** The ResNet50 model has proven its versatility across different image recognition contexts, from natural scenes and animals to technological and architectural subjects. This adaptability is key for tasks requiring a model that can handle a wide range of object types and categories.

Conclusion: Taking into account the user's recommendation, it is clear that ResNet50 stands out for its ability to provide detailed and accurate recognition across a broad array of images. Its balance of depth, precision, and versatility makes it a strong candidate for scenarios requiring detailed image analysis and recognition capabilities. The preference for ResNet50 underscores the importance of choosing a model that aligns with the specific needs and goals of the imaging task at hand, particularly where detailed outputs are paramount.

Further Addendum to Image Classification Model Analysis Report: Incorporating Ali's Perspective on Size Efficiency

Revised Recommendation Based on Size Efficiency: MobileNet

Considering the user's additional insights regarding the importance of model size for their specific use case, MobileNet emerges as a highly recommended alternative to ResNet50. The emphasis on size efficiency highlights a crucial aspect of model selection—resource constraints and deployment environments.

Key Considerations for MobileNet's Advantages:

- **Compact Architecture:** MobileNet is engineered with a focus on minimizing size and computational overhead without significantly compromising accuracy. Its use of depthwise separable convolutions effectively reduces the number of parameters, making it substantially lighter than more traditional architectures like ResNet50.
- **Deployment Flexibility:** The reduced size and efficiency of MobileNet make it particularly suitable for deployment on mobile devices and embedded systems, where memory and processing power are limited. This opens up possibilities for real-time applications and on-device processing, which are critical for many modern AI-driven services.
- **Energy Efficiency:** In addition to its compact size, MobileNet's architecture is optimized for lower energy consumption. This aspect is crucial for battery-powered devices, making MobileNet a sustainable choice for long-term deployment in mobile applications.
- **Adaptability:** Despite its streamlined architecture, MobileNet does not significantly compromise on accuracy, maintaining competitive performance across a wide range of image classification tasks. This adaptability ensures that users prioritizing size and efficiency do not have to settle for substantially lower accuracy rates.

Conclusion: Incorporating the user's perspective on the importance of model size and efficiency, MobileNet presents itself as a formidable alternative to ResNet50 for scenarios where resource constraints are a critical factor. Its balance of size, efficiency, and competitive

accuracy make it a suitable choice for applications requiring lightweight yet effective image classification solutions.
This nuanced approach to model selection underscores the importance of aligning model characteristics with specific project requirements and constraints. Whether the priority is achieving the highest level of detail and accuracy possible (as with ResNet50) or optimizing for size and efficiency for deployment on resource-constrained devices (as with MobileNet), the choice of model should be informed by a clear understanding of the task's unique demands.