Comparative analysis of classical machine learning methods and transfer learning methods using the example of classifying photos of cavitation bubbles based on alcohol content

Image classification tasks are widely used in scientific and industrial applications. Approaches to solving these tasks can be divided into two main groups: classical (traditional) machine learning methods and more modern deep learning methods, such as convolutional neural networks (CNNs) and transfer learning. Understanding the strengths and weaknesses of each approach is crucial for selecting the optimal tool for a specific task.

Classical machine learning methods. Two classical algorithms were implemented and tested: random forest (an ensemble algorithm that builds multiple decision trees and aggregates their results) and support vector machine, SVM (an algorithm that finds a hyperplane that optimally separates classes in feature space).

Unlike CNN, classical methods require manual feature extraction. For this task, features such as histogram of oriented gradients (HOG) can be used to describe the shape of bubbles, and textures (such as LBP - Local Binary Patterns) can be used to describe the structure and statistics of color channels.

Transfer learning methods. Pre-trained convolutional neural networks were used: mobilenetV3 Small (a modern, lightweight and fast architecture optimized for mobile devices) and resnet18 (a deeper architecture known for its resistance to the problem of a disappearing gradient due to the use of residual blocks). The models pre-trained on a large dataset were fine-tuned on the provided dataset with images of cavitation bubbles.

Results and analysis. The table shows summary metrics for all tested models. The metrics Precision, Recall and F1-score are presented as weighted average values to ensure a correct comparison.

Method	Accuracy	Precision	Recall	F1-score
Random forest	98.0%	0.9805	0.9801	0.9782
SVM	94.2%	0.8879	0.9423	0.9143
MobilenetV3 small	98.88%	0.9892	0.9888	0.9888
Resnet18	**99.87%**	**0.9987**	**0.9987**	**0.9987**

Random Forest performed very well, achieving an accuracy of around 98% and balanced F1 metrics. This indicates that the extracted features were of high quality and relevant to the task.

SVM, on the other hand, showed a significantly lower accuracy and, more importantly, a low Precision (0.8879) with a relatively high Recall (0.9423). This means that the SVM model often made mistakes in favor of increasing the number of positive predictions (false positives), but it also correctly identified most of the actual positive examples.

MobilenetV3 Small outperformed both classical methods, demonstrating the highest accuracy (98.88%) and perfectly balanced metrics.

Resnet18 showed the best result, approaching perfection (99.87%). All metrics for this model are almost 1.0, indicating almost perfect classification quality.

Comparing advantages and disadvantages

Criteria	Classical machine learning	Transfer learning
Performance (Accuracy)	**High (RF), but limited by features.** SVM showed significantly lower accuracy, highlighting the approach's sensitivity to the choice of algorithm and feature engineering.	**Significantly superior,** state-of-the-art. ResNet18 demonstrates near-perfect classification, a level unattainable by classical methods on this task.
Data requirements	A smaller volume of data can be sufficient for training.	Requires a sufficiently large dataset for fine-tuning, though less than training from scratch.
Computational complexity	Training is relatively fast (minutes/hours on CPU).	Training is computationally intensive (hours/days), typically requiring a GPU for efficiency. Inference can be fast, especially with optimized architectures like MobileNet.
Flexibility & Automation	**Key drawback:** Requires manual feature engineering and selection. The model's performance is highly dependent on this step, introducing a significant "human factor" and potential for suboptimal performance (as seen with SVM).	**Key Advantage:** Features are learned automatically and hierarchically, directly from the raw image data. This eliminates the need for manual feature engineering and adapts to the specific problem.
Interpretability	**Relatively high.** It is possible to analyze feature importance (e.g., in Random Forest), providing some insights into which characteristics of the bubbles were most relevant for the model's decision.	**Very low ("Black Box").** It is extremely difficult to understand which specific visual patterns in the image led to a particular classification decision, limiting model transparency.
Development time	More time is consumed by the iterative process of feature engineering, selection, and validation.	Development time is reduced due to the use of pre-trained models and a more standardized pipeline focused on fine-tuning.
Inference resources	Low, models can run efficiently on CPU.	Varies by model architecture (Mobilenet is designed for mobile/embedded devices, while

		Resnet requires more computational power).

Conclusions

1. Transfer learning methods (Resnet18, MobilenetV3) demonstrated a statistically significant advantage over classical methods. Even the most effective classical model (Random Forest) lost to the least effective CNN (MobilenetV3) by almost 1% in terms of accuracy, which can be critical in high-precision applications. The advantage of Resnet18 is overwhelming.

2. The success of transfer learning confirms that convolutional networks are capable of automatically extracting more complex, hierarchical, and relevant features than those that can be created by a human expert. This eliminates a key bottleneck and source of errors in the classical pipeline. The poor performance of SVM compared to Random Forest clearly demonstrates how fragile classical methods can be when features and algorithms are not perfectly selected.

3. For maximum accuracy without strict hardware limitations, Resnet18 should be chosen. For systems with limited resources (mobile devices, embedded systems), the ideal compromise is MobilenetV3, which outperforms classical methods while remaining lightweight. In situations where interpretability of solutions is critical (e.g., for scientific understanding of a phenomenon), a well-tuned Random Forest may be preferable, as it allows for analyzing the importance of different features.

4. The outstanding results of all models (except for SVM) confirm that the alcohol content in the solution creates clearly distinguishable visual patterns in the images of cavitation bubbles. However, CNNs are capable of capturing the more subtle and complex nuances of these patterns.

For the task of classifying cavitation bubble images based on alcohol content, transfer learning methods are the optimal choice, providing unparalleled accuracy and reliability. While the classic Random Forest method has shown excellent results and can be used in scenarios that require interpretability, its accuracy is limited by the quality of manual feature extraction. For industrial or scientific applications, ResNet18 can be used due to its high accuracy. If computational resources are limited, MobilenetV3 Small serves as an excellent alternative with high accuracy and efficiency.