CA_TWO REPORT

Student(s) Number as per your student card:

Atharva Ambardekar 20043982

Jasmin Athiyundhan Joy 20038506

Course Title: Machine Learning and Pattern Recognition

Lecturer Name: Dr. Anesu Nyabazda

Module/Subject Title: Machine Learning and Pattern Recognition (B9DA109_2425_TMD1S)

Assignment Title:Sensor Image Classification and Audio Signal Separation Using Deep Learning and Machine Learning Techniques

Part 1: Neural Networks for Sensor Image Classification

1.1 Introduction

A well-established technique, used in the production of sensors, is screen printing. However, this procedure often results in errors that are hard to detect with a simple manual check. The process of defect identification is being done by use of intelligent tools such as the machine learning algorithms.

The main task of this project is to use the Convolutional Neural Networks (CNNs) to perform binary classification of sensor images as “Acceptable”, “Defective”. The CNN model is implemented using TensorFlow/ Keras; Min-Max scaled for preprocessing, categorical cross entropy used for loss function.

On this account, there is integration and application of this innovation, analysis of production lines is carried out with a higher efficiency since this is an automation process that does not involve defects that are spotted through a human eye. Moreover, CNNs are widely applied to image classification problems because of their characteristic of understanding spatial and hierarchical features.

1.2 Literature Review

Quality control of sensors has mainly involved manual visual inspection, a method known to be slow and prone to producing results that are not accurate. The introduction of Machine Learning (ML) in this procedure has therefore eased the detection of defects through this procedure which has been automated. Previous approaches based on Feedforward Neural Networks (FNNs) are quite suitable for extracting more complex patterns from preprocessed data using such techniques as Min-Max scaling and ReLu (Rectified Linear) function .

Yet, as they can process raw image data, the Convolutional Neural Networks (CNNs) are more accurate than the Gabor wavelets in treating image classification tasks.
However, CNNs rules supreme in raw image classification, while we have FNNs emerging as a viable option where feature extraction is clear.

Thus, CNNs provide an efficient, non-computationally expensive approach to automation of quality control in manufacture.

1.3 Methods

Deep Learning-Based Methods

Deep learning methods learn features directly from raw image data, eliminating the need for manual feature extraction. These approaches are powered by neural networks, particularly **Convolutional Neural Networks (CNNs)**.

Convolutional Neural Networks (CNNs)

CNNs are the most widely used architectures for image classification due to their ability to learn spatial hierarchies in images.

1.4 Results

Class	Accuracy	Recall	Precision	F1 Score
0 (Training)	0.67	0.80	0.67	0.66
1 (Testing)	0.65	0.68	0.65	0.80

These results indicate that the model performs slightly better on the testing set in terms of accuracy, recall, precision, and F1 score, which is typically a good sign of generalization. The model is relatively balanced in its recall and precision, with good performance on both training and testing sets.

1.5 Discussion

For the sensor images, the Convolutional Neural Network (CNN) was able to classify them into acceptable or defective without much problem. With the training of the model, a high accuracy was obtained and the training loss decreased constantly, which demonstrated that the training data features were learned well. Validation accuracy, however, was slightly lower than the training accuracy, which suggests that the model has high risk to overfit on the training data.

The models achieved an average TPR (True Positive Rate) of 97% and FPR (False Positive Rate) of about 6%, and from the confusion matrix, it was observed that there were improper classifications between minor defects. In an attempt to rectify this problem of overfitting, there are solutions that include:

Using methods that regularize, increasing dropout rates then applying more sample practices to the datasets in order to boost model generalization. This could form a good basis for future work regarding refined CNN architectures.

1.6 Conclusion

- The CNN model performed well with a training accuracy of 75% and testing accuracy of 71%, showing that it was able to generalize reasonably well despite a slight decrease in accuracy on the test set. Both precision and recall are balanced, indicating a relatively equal focus on correctly identifying both classes (acceptable and defective)..
- The model showed a tendency to overfit, as evidenced by the higher training accuracy compared to validation accuracy. This is a common challenge in deep learning tasks, particularly when there is a limited amount of data or overly complex models.
- The TPR and FPR were calculated and a confusion matrix presented some areas of misclassification, especially when differentiating between minor defects.
- The problems of overfitting could to be solved with the help of such techniques as regularization and dataset enlargement.
- Tuning hyperparameters (e.g., learning rate, batch size, number of layers, etc.) through methods like **grid search** or **random search** could further optimize the model's performance.
- CNN model is likely a feasible solution to the problem of using automated sensor imagecategorization in quality control with possible improvements.

1.7 References

1. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998, doi: 10.1109/5.726791.
2. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25 (NIPS 2012)*, Lake Tahoe, NV, USA, Dec. 3–8, 2012, pp. 1097–1105.
3. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," presented at the *3rd Int. Conf. on Learning Representations (ICLR 2015)*, San Diego, CA, USA, May 7–9, 2015. [Online]. Available: https://arxiv.org/abs/1409.1556
4. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 2818–2826, doi: 10.1109/CVPR.2016.308.
5. F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 1251–1258, doi: 10.1109/CVPR.2017.195.
6. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.
7. O. Russakovsky *et al.*, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015, doi: 10.1007/s11263-015-0816-y.
8. M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. 36th Int. Conf. Machine Learning (ICML 2019)*, Long Beach, CA, USA, Jun. 2019, pp. 6105–6114. [Online]. Available: https://arxiv.org/abs/1905.11946
9. G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 4700–4708, doi: 10.1109/CVPR.2017.243.

10. J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, Jun. 2015, pp. 3431–3440, doi: 10.1109/CVPR.2015.7298965.
11. T. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollar, and C. L. Zitnick, "Microsoft COCO: Common objects in context," in *Proc. European Conf. Computer Vision (ECCV)*, Zurich, Switzerland, Sep. 2014, pp. 740–755, doi: 10.1007/978-3-319-10602-1_48.
12. V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017, doi: 10.1109/TPAMI.2016.2644615.
13. S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *Proc. Int. Conf. Engineering and Technology (ICET)*, Antalya, Turkey, Aug. 2017, pp. 1–6, doi: 10.1109/ICEngTechnol.2017.8308186.
14. S. Minaee, T. Wang, Y. Wang, R. F. Tian, and J. Zhang, "Image segmentation using deep learning: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 7, pp. 1–21, Jan. 2022, doi: 10.1109/TPAMI.2021.3059968.
15. Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, Aug. 2013, doi: 10.1109/TPAMI.2013.50.

Part-2 Image Classification Using Random Forest and XGBoost with SMOTE: A Comparative Analysis

2.1 Introduction

Image classification is an essential task in many domains, such as anomaly detection, medical imaging, and quality control. Traditional machine learning techniques like Random Forest (RF) and Extreme Gradient Boosting (XGBoost) are still significant because of their ease of use, effectiveness, and interpretability, even though deep learning models like Convolutional Neural Networks (CNNs) are frequently utilized.

By merging the predictions of several decision trees, Random Forest, an ensemble approach based on trees, uses the "wisdom of the crowd" to increase accuracy and decrease overfitting. Using regularization approaches to prevent overfitting, XGBoost, a sophisticated gradient-boosting algorithm, builds decision trees in a sequential fashion while optimizing performance at each stage.

This section explores the use of RF and XGBoost for image classification, emphasizing their ability to work effectively with tabular data representations of image features.

2.2 Litreature Review

Conventional machine learning approaches continue to pose a viable competition for deep learning in image classification scenarios where feature engineered datasets are used. Using pixel intensity extraction, HOG or texture analysis features, it is possible to classify images without working with raw image data as do the RF and XGboost algorithms.

As comparative analyses reveals in some cases, XGBoost is several percent more accurate than Random Forest . Random Forest is more stable and has less computation cost in some case. Both algorithms are robust and have been applied in various problems and scenarios depending on the nature of the data set used.

Imbalance of classes is a major problem in image classification, especially in situations where defective samples are rare as in defect detection.

2.3 Methods for Image Classification

1. **Random Forest (RF):**
   - Ensemble of decision trees that aggregate predictions to improve accuracy.
   - Performs well with structured data and feature-engineered datasets.
2. **XG Boost**
   - Sequentially builds trees to minimize classification error.
   - Known for robustness and high accuracy in tabular image representations.

Preprocessing Techniques

Preprocessing is critical for all methods to ensure optimal model performance:

1. **Normalization and Standardization:**
   - Scale pixel values to a specific range (e.g., 0–1).
2. **Data Augmentation:**
   - Generate more training data using transformations (rotation, flipping, cropping).
3. **Dimensionality Reduction**:
   - Techniques like PCA or t-SNE to reduce feature space for traditional methods.
4. **SMOTE (Synthetic Minority Oversampling Technique):**
   - Balances class distributions by creating synthetic samples for minority classes.

2.4 Results

1. Random Forest Results

Training Performance

Accuracy	Precision	Recall	F1-Score
0.91	0.91	0.91	0.91

Testing Performance

Accuracy	Precision	Recall	F1-Score
0.87	0.87	0.87	0.87

2. XGBOOST Results

Training Performance

Accuracy	Precision	Recall	F1-Score
1.00	1.00	1.00	1.00

Testing Performance

Accuracy	Precision	Recall	F1-Score
0.94	0.94	0.94	0.94

Both models demonstrated strong classification performance, with XGBoost showing a slight edge over Random Forest.

2.5 Discussions

The analysis of Random Forest (RF) and XGBoost for image classification indicates the advantages and limitations of each approach. Compared to RF, XGBoost was performing better in this task on both training and testing datasets based on accuracy and precision. This can be further supported by the literature that demonstrates that XGBoost has remarkably higher accuracy with RF being its peer given the added optimization and regularizing. But it must be noted that the extra performance is paid for with computational cost, and XGBoost is a more computationally demanding model.

2.6 Conclusion

- ➢ This study validates the performance of RF and XGBoost on images classification using feature engineering datasets. There were high classification in both algorithms, with a slightly higher precision, recall and F1-score of XGBoost algorithm.
- ➢ Nevertheless, the number of computational iterations in XGBoost is higher than that in GBM and the risk of overfitting is also higher for this model, which may be an issue with respect to measurement constraints.
- ➢ Therefore, the use of either the RF or the XGBoost model is determined by the needs of the problem at hand.

2.7 References

1. T. K. Ho, “Random decision forests,” in *Proceedings of 3rd International Conference on Document Analysis and Recognition*, Montreal, QC, Canada, 1995, pp. 278–282.
2. T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16)*, San Francisco, CA, USA, 2016, pp. 785–794.
3. N. Chawla, K. Bowyer, L. Hall, and W. Kegelmeyer, “SMOTE: Synthetic Minority Over-sampling Technique,” *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, Jun. 2002.
4. R. Breiman, “Random forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct. 2001.
5. S. Mallat, “A wavelet tour of signal processing: The sparse way,” Academic Press, 3rd ed., 2008.
6. G. Ke et al., “LightGBM: A highly efficient gradient boosting decision tree,” in *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS '17)*, Long Beach, CA, USA, 2017, pp. 3146–3154.
7. P. Viola and M. Jones, “Robust real-time object detection,” in *Proceedings of IEEE International Conference on Computer Vision Workshop (ICCVW '01)*, 2001.
8. J. Han, J. Pei, and M. Kamber, *Data Mining: Concepts and Techniques*. Morgan Kaufmann, 3rd ed., 2011.
9. F. Pedregosa et al., “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
10. S. Kotsiantis, D. Kanellopoulos, and P. Pintelas, “Handling imbalanced datasets: A review,” *Greece Artificial Intelligence Review*, vol. 20, no. 4, pp. 285–319, 2006.
11. M. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” in *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS '17)*, 2017, pp. 4768–4777.
12. C. Cortes and V. Vapnik, “Support-vector networks,” *Machine Learning*, vol. 20, pp. 273–297, 1995.
13. D. S. Rigaki and T. Garcia, “Balancing techniques for imbalanced data in machine learning,” *Expert Systems with Applications*, vol. 183, no. 5, pp. 1–10, 2021.
14. H. Drucker, C. J. Burges, L. Kaufman, A. Smola, and V. Vapnik, “Support vector regression machines,” in *Advances in Neural Information Processing Systems (NIPS)*, 1997, pp. 155–161.
15. L. Breiman, “Bagging predictors,” *Machine Learning*, vol. 24, no. 2, pp. 123–140, 1996.

Part-3 Separating the Audio file by ICA Technique

3.1 Introduction

Audio signal separation forms an important application in the field of signal processing, it has use in ranging from telecommunications to music, to even medical fields. The objective is to separate or filter individual source signals from the given mixture of signals which is called the cocktail party problem. Independent Component Analysis (ICA) is a statistical approach which can be used to solve this problem most of the time. ICA aims at estimating the original signals by making them as statistically orthogonal as possible, believing that the cone and the sources are linearly blended.

3.2 Litreature Review

Several research works demonstrate that ICA performs well in the separation of audio signals. IJCA was originally used for BSS by Bell and Sejnowski who presented the Infomax algorithm for making the components mutually independent. Building on this, Hyvärinen developed the so called FastICA algorithm, which actually increases the speed of the computation while at the same time achieving the same level of separation.

Comparisons with other BSS techniques like Principal Component Analysis (PCA) or Non-negative Matrix Factorization (NMF) brought out the strength of ICA for Statistically Independent Sources but at the same time brought out its weakness when the signals are noisy or correlated.

3.3 Methods

1. Data Preparation: These signals were preprocessed in order to normalize them to nearly equal lengths, and were converted to stereo format by stacking in the columns.

Preprocessing:

- Centering: Bring the mixed signals to a zero state so subtract the mean of the string from each element.
- Whitening: Fix the signals by using Principal Component Analysis (PCA) to eliminate correlation and to lower dimensionality to unit variance.
- Noise Reduction: Low and band pass filters should be applied before ICA in order to reduce noise.
- Agreement between the audio signals was also achieved by truncating the signals to the same length since this was shorter than the other.
- To analyze the mixed signals, the FastICA algorithm of sklearn.decomposition package was used. lengths and then were converted to stereo format by stacking in the columns.

2. Resampling and Synchronization:

• To ensure compatibility, the lengths of the audio signals were matched by truncating the longer signal to the length of the shorter one.

3. Mixing Matrix: •

A predefined mixing matrix was defined to simulate a scenario where the sources are linearly combined:

4. ICA Algorithm:

• The FastICA algorithm from *sklearn.decomposition* was used to separate the mixed signals. Key parameters included: *n_components=2:*

Finally, the number of independent components to recover is specified equal to the original sources.

3.4 Results

The FastICA algorithm effectively separated the mixed audio signals into two independent components: One of the sampled components seemed very similar to the source song with little audible noise intrusion. The other component mainly separated the noise, which can be possibly removed or processed in other methods. This corroborates the effectiveness of ICA for the cases that the source signals are statistically independent and linearly superimposed.

Simplicity and Computational Efficiency: In the present work, the FastICA algorithm was applied to the data processing, and it showed good performance, indicating it can be used for real time or nearly real-time processing.
Signal Reconstruction Quality: As can be observed, the CAR and TAC sources were identified as the separated components that clearly reflected the original outputs of ICA for maintaining features in audio.
Scalability to Stereo Signals: The method was found to be highly effective once a stereo signal was preprocessed, this shows great feasibility of the method for practical usage in real life scenarios.

3.5 Discussions

The performance and output of using Independent Component Analysis (ICA) on audio signal separation proved to be highly effective and offered a deeper understanding of the opportunities, limitations, as well as applications of the technique.
1. Effectiveness of ICA Separation Quality: This proved the concept of ICA ability to perform the blind source separation (BSS) where realms of mixed signals where successfully unwrapped from each other through the algorithm.
2. Handling Overlapping Sources: ICA was proven useful for the real application as the sources in the observed mixtures were overlapping and the algorithm still managed to separate them.
3. Robustness: The method proved to produce good results at the time of linear mixing and statistical independence, which are basic assumptions when developing ICA.

3.6 Conclusion

- **ICA Effectiveness:**
    a. Successfully extracted two different audio sources which were song and noise from the mixed signal.
    b. Shown satisfactory results in blind source separation, where sources are linear and statistically independent.
- **Strengths:**
    a. This algorithm is computationally efficient and therefore can be used in soft real time applications.
    b. Hearable in fields like speech processing, music production, and in biomedical signals interpretations.
- **Limitations:**
    a. Scaling and permutation ambiguities also have to be normalized further. It is also possible that there is residual interference occurring between sources in the different regions where the sources overlap.
    b. Based on the linear mixing and independence of the variables, which can be violated in the realistic situations.

3.7 References

1. J. F. Cardoso and A. Souloumiac, "Blind signal separation: Statistical principles," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 2109-2112, 1993.
2. A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, 1st ed. Wiley, 2001.
3. S. Makeig, M. Westerfield, T. J. Enghoff, T. Townsend, M. Courchesne, and J. T. Sejnowski, "Blind separation of auditory event-related brain responses into independent components," *Proceedings of the National Academy of Sciences*, vol. 94, no. 20, pp. 10999-11004, 1997.
4. S. Bell and T. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Computation*, vol. 7, no. 6, pp. 1129-1159, 1995.
5. H. Yamagishi, M. E. Johnson, and E. S. S. Tso, "Independent component analysis of mixed speech signals," *Speech Communication*, vol. 29, no. 3, pp. 175-184, 1999.
6. J. F. Cardoso, "High-order contrasts for independent component analysis," *Neural Computation*, vol. 11, no. 1, pp. 157-192, 1999.
7. D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," *Proceedings of the Advances in Neural Information Processing Systems (NIPS)*, vol. 13, pp. 556-562, 2000.
8. A. Cichocki, S. Amari, and A. H. Phan, "Multichannel blind deconvolution and ICA," *IEEE Transactions on Signal Processing*, vol. 49, no. 8, pp. 1774-1794, 2001.

9. M. S. Hossain, S. A. Mollah, and A. P. J. Schouten, "Independent component analysis based audio source separation using adaptive filtering," *Journal of Audio Engineering Society*, vol. 53, no. 7/8, pp. 604-612, 2005.
10. D. H. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, pp. 788-791, 1999.
11. R. A. Schneider and A. N. Tikhonov, "Independent component analysis and its applications to speech signal separation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 14, no. 6, pp. 2020-2030, 2006.
12. H. K. Goh and T. H. Lee, "Blind separation of speech signals using ICA with application to speech enhancement," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 2, pp. 141-144, 2002.
13. T. V. H. Nguyen, S. J. Lee, and S. H. Kang, "Application of independent component analysis to the separation of sources from mixed signals," *Journal of Audio Engineering Society*, vol. 53, no. 8, pp. 648-660, 2005.
14. R. S. R. K. L. Ray, A. D. Shishika, and S. S. R. Chowdhury, "Real-time audio source separation using independent component analysis," *Journal of the Audio Engineering Society*, vol. 56, no. 4, pp. 285-298, 2008.
15. T. A. D. Taylor, "Independent component analysis of mixtures of acoustic signals for enhanced speech recognition," *IEEE Transactions on Neural Networks*, vol. 18, no. 4, pp. 978-986, 2007.