# CW Paper Summaries 16/10/2024

## Marked Up Papers summarised by ChatGPT

@inproceedings{song2022apple,

title={Apple Counting Network Before Fruit Thinning Period Based On Dilated Convolution},

author={Song, Mengyang and Jiang, Guoquan and Huo, Zhanqiang and Yang, Zhengyuan and Zhang, Hongxu},

booktitle={Proceedings of the 2022 11th International Conference on Networks, Communication and Computing},

pages={129--134},

year={2022}

}

The paper **“Apple Counting Network Before Fruit Thinning Period Based on Dilated Convolution”** aims to develop an efficient, automated system for counting apples before the thinning period, aiding precision agriculture. This is crucial because accurate fruit counting informs yield estimation, which is essential for orchard management. Traditional methods are manual and labor-intensive, and the paper proposes an alternative using deep learning.

The experiment employs **FTACNet**, a novel model based on dilated convolution integrated with a **Bayesian loss function**. Dilated convolution, which expands the receptive field without losing resolution, helps handle challenges like occlusion and varying fruit sizes. The model uses the **VGG19** backbone for feature extraction, combined with dilated convolution layers to process images captured via smartphones. Images are annotated using point labeling, and the network generates density maps from which apple counts are estimated.

The goal is achieved successfully, with **FTACNet** outperforming existing methods like CSRNet in terms of mean absolute error (MAE) and mean square error (MSE), achieving 4.14 MAE and 5.62 MSE. The model's small size makes it deployable on mobile devices, demonstrating its potential for real-world orchard applications.

@article{ye2024vision,

title={Vision foundation model for agricultural applications with efficient layer aggregation network},

author={Ye, Jianxiong and Yu, Zhenghong and Lin, Jiewu and Li, Hongyuan and Lin, Lisheng},

journal={Expert Systems with Applications},

volume={257},

pages={124972},

year={2024},

publisher={Elsevier}

}

The paper **“Vision Foundation Model for Agricultural Applications with Efficient Layer Aggregation Network”** aims to address the challenges of plant detection and counting in agricultural environments. It proposes a new model, **TasselELANet**, specifically designed to handle the complexity of real-world agricultural tasks, such as detecting maize tassels, wheat ears, and rice panicles, with high accuracy and efficiency.

The core technology used in this research is the **Efficient Layer Aggregation Network (ELAN)**, integrated with a 16-fold downsampling encoder and a decoder utilizing only two feature layers. This architecture minimizes feature redundancy while maintaining a robust gradient propagation path. The model uses **convolutional neural networks (CNNs)** optimized for agricultural vision tasks, and employs **Bayesian loss functions** and **CARAFE** upsampling to handle dense crops and occlusions. The model is evaluated on three public datasets (maize, wheat, and rice), showing superior performance compared to state-of-the-art methods like **Faster R-CNN** and **CenterNet**.

The paper successfully achieves its goals, with TasselELANet outperforming existing models in terms of accuracy (0.865 AP50) and computational efficiency. However, the model struggles with occlusions in dense agricultural environments.

# CW Paper Summaries 07/10/2024

## Marked Up Papers summarised by ChatGPT

@article{chitradevi2014overview,

title={An overview on image processing techniques},

author={Chitradevi, B and Srimathi, P},

journal={International Journal of Innovative Research in Computer and Communication Engineering},

volume={2},

number={11},

pages={6466--6472},

year={2014},

publisher={Citeseer}

}

The paper titled **“An Overview on Image Processing Techniques”** aims to provide a detailed review of various methods used in digital image processing and their broad range of applications. The authors seek to explain how techniques such as image acquisition, enhancement, segmentation, feature extraction, and classification are crucial for improving the visual quality of images for both human interpretation and machine perception. These techniques are widely applied in fields like remote sensing, medical imaging, and forensic analysis.

The technologies used include specific algorithms for enhancing and processing digital images. In image enhancement, contrast stretching is employed to expand a narrow range of pixel intensity values, while noise filtering methods, such as low-pass and median filters, are used to reduce unwanted noise in the image. Histogram modification, specifically histogram equalization, is applied to adjust the image contrast by redistributing pixel intensity values. For image segmentation, the paper discusses methods like global thresholding, which assigns a binary value to pixels based on intensity, and watershed segmentation, which identifies regions within the image using topographic gradients.

The paper successfully meets its goal by offering a comprehensive overview of image processing techniques. However, it lacks experimental data to evaluate the performance of the methods discussed. For further investigation, you may explore the work of **L. Gagnon and A. Jouan** on “Speckle Filtering of SAR Images” for a deeper understanding of wavelet-based filtering techniques​.

@article{xiao2024apple,

title={Apple ripeness identification from digital images using transformers},

author={Xiao, Bingjie and Nguyen, Minh and Yan, Wei Qi},

journal={Multimedia Tools and Applications},

volume={83},

number={3},

pages={7811--7825},

year={2024},

publisher={Springer}

}

The paper **“Apple ripeness identification from digital images using transformers”** aims to develop a non-destructive method for determining apple ripeness using digital images and deep learning models. The goal is to automate the classification of fruit ripeness, which could aid in agriculture by enhancing fruit-picking automation, especially in environments like New Zealand where labor shortages exist. The paper primarily focuses on comparing two object detection methods—YOLO (You Only Look Once) and the Transformer model—to classify the ripeness of apples and pears.

The experiment uses YOLOv5 and the Transformer model as the core technologies. YOLOv5 works by dividing an image into grids and predicting bounding boxes for objects within each grid, using parameters like object size and confidence score. The Transformer model, originally designed for natural language processing, is employed to process images by converting them into vectorized sequences and applying self-attention mechanisms to identify fruit categories and maturity levels. The experiments were conducted using a dataset of 2,000 images taken with mobile cameras, labeled to classify fruit as either ripe, overripe, or of other categories.

The paper successfully achieves its goal by showing that YOLOv5 outperforms the Transformer model in both detection speed and accuracy. A follow-up reference is **Bochkovskiy et al. (2020)** on YOLOv4, which optimizes both speed and accuracy in object detection​.

@article{salvi2021impact,

title={The impact of pre-and post-image processing techniques on deep learning frameworks: A comprehensive review for digital pathology image analysis},

author={Salvi, Massimo and Acharya, U Rajendra and Molinari, Filippo and Meiburger, Kristen M},

journal={Computers in Biology and Medicine},

volume={128},

pages={104129},

year={2021},

publisher={Elsevier}

}

The paper **“The Impact of Pre- and Post-Processing on Deep Learning Frameworks for Digital Pathology Image Analysis”** reviews how pre- and post-processing techniques are integrated into deep learning models to enhance performance in medical image analysis, particularly in digital pathology. The authors aim to demonstrate how these techniques improve tasks like classification, detection, and segmentation, which are essential in analyzing histology images for cancer detection, cell counting, and more.

In their experiments, the authors utilize Convolutional Neural Networks (CNNs) as the core technology. Pre-processing involves methods such as stain normalization to standardize image color variations and tissue artifact detection to remove noise like tissue folds and blurring. Techniques like global thresholding, histogram specification, and k-means clustering are applied to optimize image quality. Post-processing techniques such as connected component analysis and non-maxima suppression refine the segmentation and classification outcomes by correcting prediction errors and excluding false positives.

The paper achieves its goal by presenting a detailed overview of these techniques and demonstrating their impact on increasing accuracy in digital pathology. However, the review lacks experimental comparison to alternative machine learning models.

For further investigation, explore **Litjens et al. (2017)**, which discusses automated CNN-based pathology classification techniques​