Deep Learning methods for biotic and abiotic stresses detection in fruits and vegetables: state of the art and perspectives

S. C. Ariane Houetohossou[1], Ratheil V. Houndji[1,2], Castro G. Hounmenou[1], Rachidatou Sikirou[3], and Romain L. Glèlè Kakaï[1]

[1] Laboratoire de Biomathématiques et d'Estimations Forestières, Faculty of Agronomic Sciences, University of Abomey-Calavi, 04 BP 1525 Cotonou, Republic of Benin
[2] Institut de Formation et de Recherche en Informatique, University of Abomey-Calavi, 01 BP 526 Cotonou, Republic of Benin
[3] Laboratoire de Défense des Cultures, Centre de Recherches Agricoles d'Agonkanmey, Institut National des Recherches Agricoles du Bénin (INRAB), 01 BP 884 Cotonou, Republic of Benin

Abstract

- Deep Learning (DL), a type of Machine Learning, has gained significant interest in many fields, including agriculture.
- ResNet50 and VGG16 were the most used architectures from the 132 reviewed articles.
- Data scarcity, imbalance, and homogeneity of some image backgrounds, which negatively influence the robustness of the developed models were discussed.

Introduction

- Fruits and vegetables contain dietary fiber and vitamins, which help lower the risk of cardiovascular disease and obesity (Slavin, 2012).
- However, many biotic and abiotic factors cause losses in their productivity.
- Deep Learning (DL) is use for early disease identification.
- We performed a bibliometric analysis and a systematic literature review focusing on the two types of stresses for effective monitoring to enhance crop performance.

Methodology

Paper selction

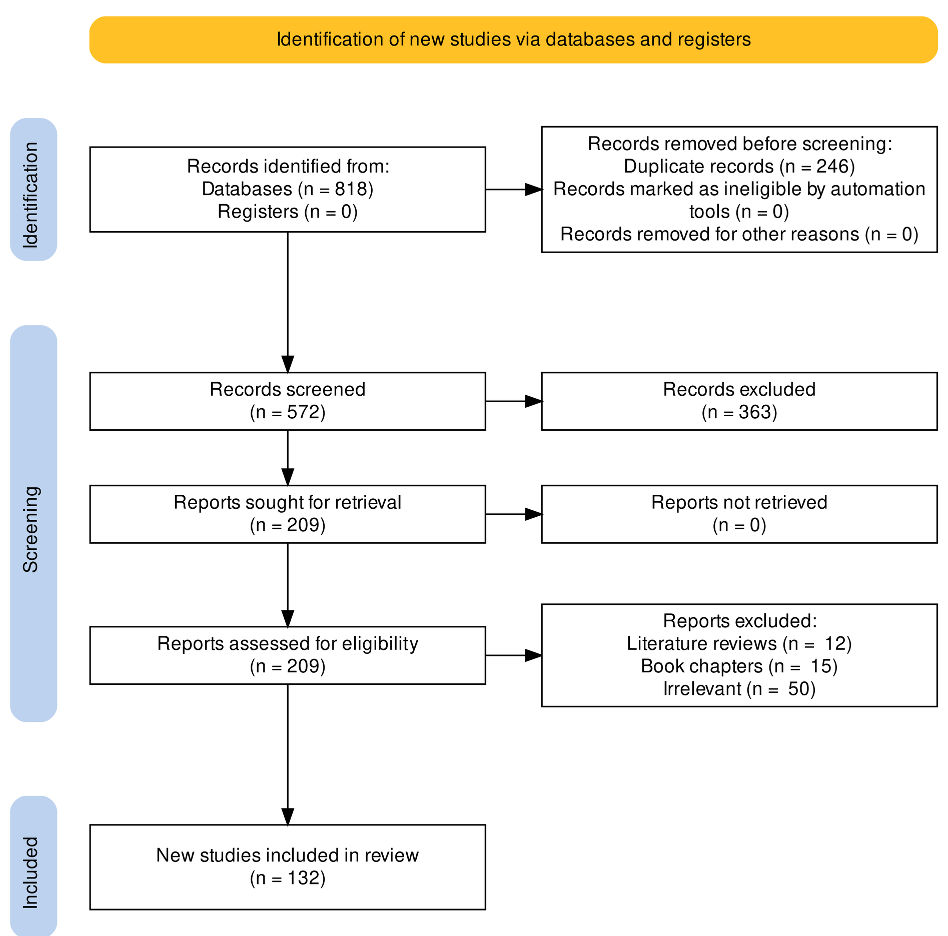

Figure 1: PRISMA flow diagram of the selection process

Research Questions (RQ)

We state below our research questions:

1. What is the motivation for using the DL or AI method?
2. What species was concerned?
3. What type of stress was involved?
4. What are the types and sources of data used?
5. What are the countries of the self-made data?
6. What models were used?
7. What are the evaluation metrics?
8. What are the performances achieved?
9. What are the gaps and perspectives?

Tools used for the literature synthesis and analysis

- VOSviewer for keywords co-occurrence network (Eck, 2022)
- Pandas, Matplotlib, and Numpy libraries of Spyder Notebook in Anaconda environment and package 'ggplot2'of R software.

Main Results

General statistics

- 132 articles published between 2003 and 2022
- Journal articles (64%)
- Conference proceedings (36%)

Bibliometric analysis

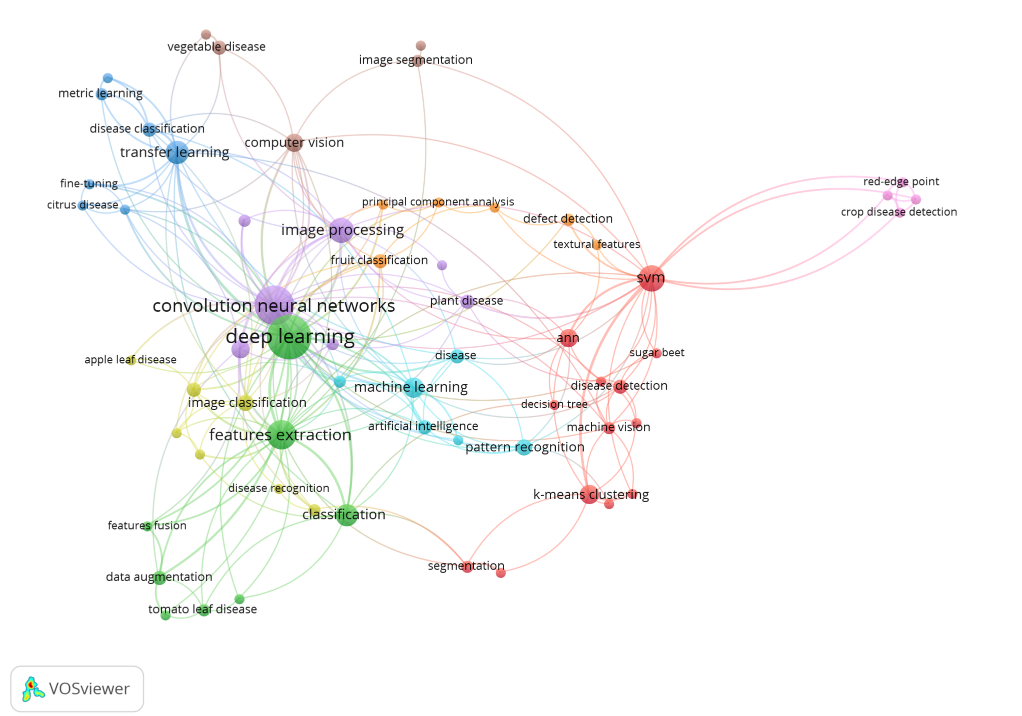

Figure 2: Keywords co-occurrence

RQ1: What is the motivation for using the DL or AI method?

- Biotic stress detection (diseases and pest attacks): 67.72%,
- Abiotic stress detection (nutrient deficiencies, heavy metal contamination, water stress, light stress): 5.51%,
- Model improvement with complex backgrounds, increasing of speed, reduction of computation time through transfer learning: 26.77%

RQ2: What species was concerned?

Table 1: Top five most studied species of fruits and vegetables

Species	Scientific name	Occurrence	Frequency (%)
Apple	*Malus domestica*	35	16.67
Tomato	*Solanum lycopersicum*	33	15.71
Grape	*Vitis vinifera*	16	7.62
Lemon	*Citrus lemon*	14	6.67
Peach	*Prunus persica*	13	6.19

RQ3: What is the type of stress? What are the type and source of the data?

Type of stress

- Biotic stress: 93.65%
- Abiotic stress: 6.35%

Type of data

- Images: 93%,
- Climate data 7%

Source of data

- Plant Village: 29%,
- Self build: 52.15%
- Other online source: 18.85%,

RQ5: What are the countries of the self made-data?

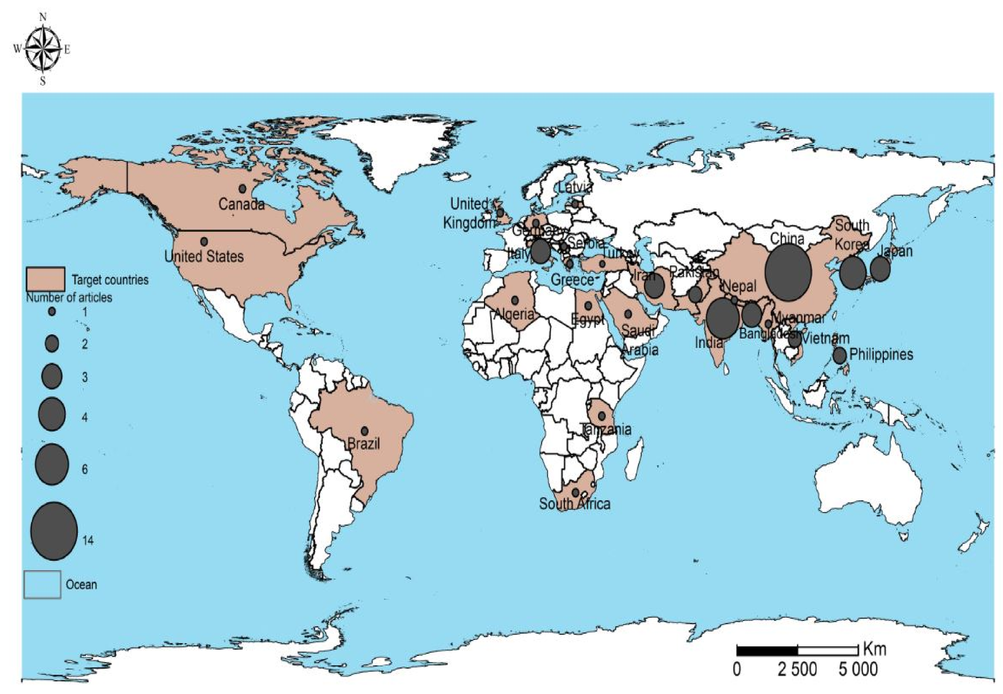

Figure 3: Countries-wise distribution of self data-collected

RQ6:What models were used?

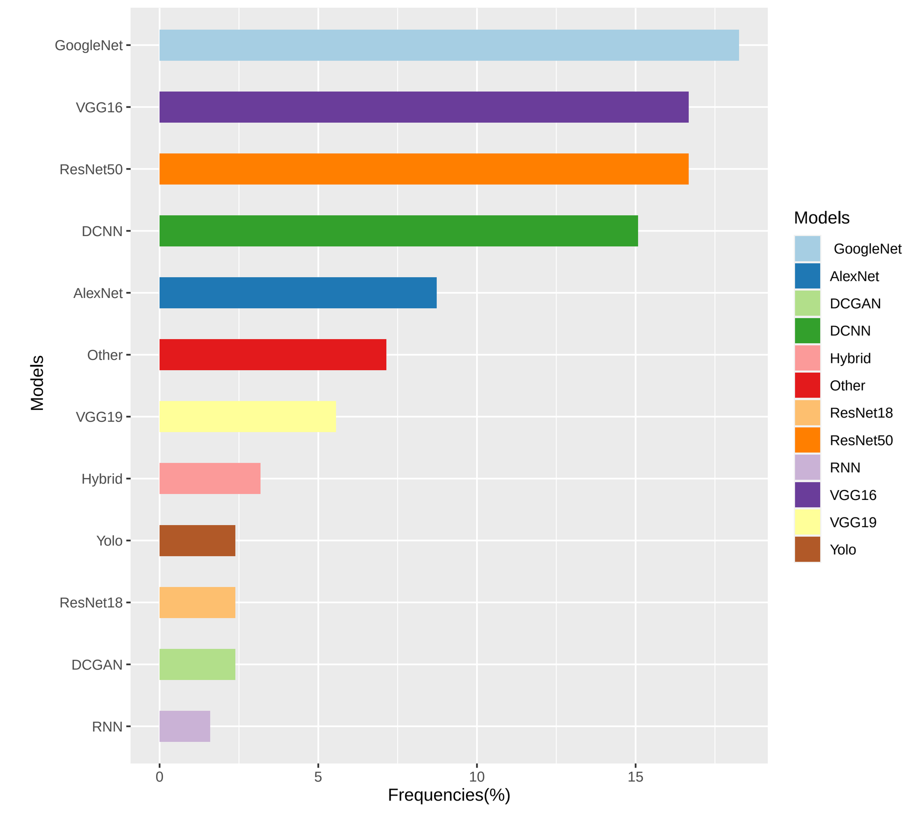

Figure 4: DL models

RQ7: What are the evaluation metrics?

Top 3 most used metrics

- Accuracy: 40%,
- Precision: 15.56%
- Recall: 10.16%

RQ8: What are the performances achieved?

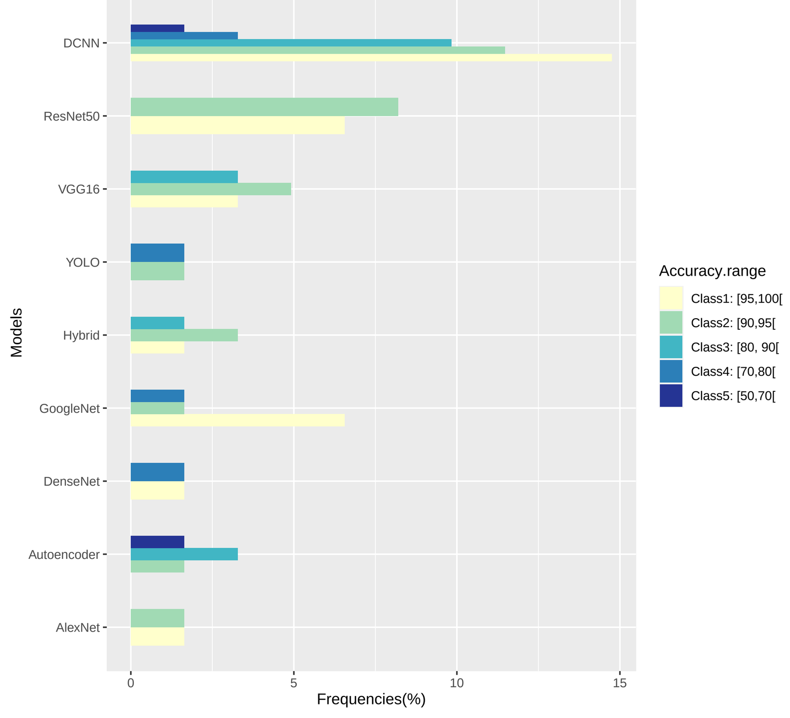

Figure 5: Performance of best models

RQ9: What are the gaps and perspectives?

Gaps

- Small database
- Unbalance class
- Homogeneous background
- Non-robustness of models
- Self collected data are mostly from developed countries
- No study on climate change stress prediction on fruits and vegetables using DL or AI

Perspectives

- Collection of data on real field situations over the world
- Improvement of the robustness of the models
- Study of the prediction of stress due to climate change on fruits and vegetables using AI and DL.

Conclusions

Despite being widely used for diseases and stress classification, DL models present many challenges for users and scientists. For better productivity of fruits and vegetables, automatic methods based on AI and DL for early identification of stress need to be improved.

References

[1] Joanne L Slavin and Beate Lloyd. Health Bene fi ts of Fruits and Vegetables 1. pages 506–516, 2012.

[2] Nees Jan van Eck and Ludo Waltman. Vosviewer manual, 2022.