TraceFinder: Forensic Scanner Identification

1. Introduction

TraceFinder is a forensic research initiative focused on pinpointing which specific scanner produced a given digital document. Although imperceptible to the naked eye, each scanner embeds subtle noise patterns and textural artefacts that function as its unique “fingerprint.” By leveraging advanced image-analysis methods and machine-learning models, the project seeks to train a system capable of learning these device-specific traits. The ultimate objective is to build a tool that can reliably identify the scanner model solely from the characteristics of the scanned image.

Milestone 1 – Dataset Collection & Preprocessing

2.1. Introduction

This phase centered on exploring the raw scanned-document dataset, examining its fundamental characteristics, and establishing a reliable preprocessing workflow. The main objective was to clean, standardize, and organize the data, ensuring it is fully prepared for feature extraction and model development in the upcoming steps.

2.2 Scanner Models Present in the Dataset

The dataset includes 11 scanner models, each contributing Flatfield, Official, and Wikipedia image samples:

The full project dataset, containing samples from multiple scanner devices, was categorized into three main labels based on content type:

- Wikipedia
- Official

2.3. Dataset Exploration and Key Observations

- An initial examination of the complete dataset—consisting of 4,590 images—highlighted several important patterns and constraints for downstream modeling. The distribution of samples was heavily skewed, with the *Wikipedia* class containing 2,368 images, while the *Flatfield* class had only 22, making the dataset notably imbalanced.
  Image resolutions were also inconsistent, clustering around two main native formats: a 150 DPI group (≈1240 × 1754 or 1236 × 1754) and a 300 DPI group (≈2480 × 3508 or 2400 × 3500). All images were captured in RGB.

- Content characteristics varied as well—Flatfield samples exhibited extremely low entropy, whereas Official and Wikipedia images showed higher entropy due to the presence of text.

- Finally, dataset organization was completed by generating manifest files (such as *flatfield_preprocessing.csv*) that map each file's full path to its corresponding scanner ID and a numerical label, ensuring clean and traceable indexing for later stages.

2.4. Preprocessing Pipeline

- A consistent preprocessing strategy was designed and applied to all images to ensure the final output is standardized and focused on noise- based features. The overall target size was set to 512 x 512. This pipeline began with Grayscale Conversion, which simplifies processing and focuses analysis on structural and noise-based features by converting all images to 8-bit grayscale (1 channel).

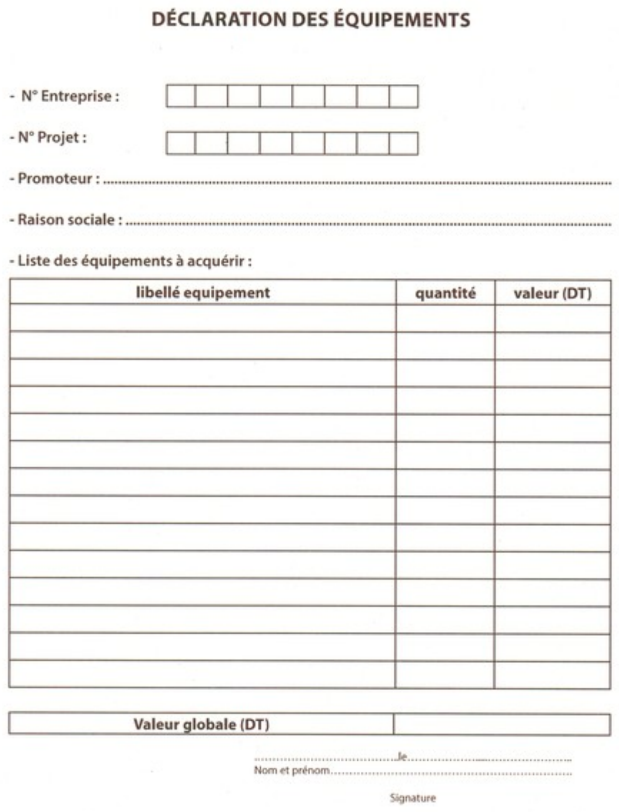
- Next, Resizing was performed to create a uniform input size required for the classifier by adjusting all images to the target dimension of 512 x 512.

- The final step was Normalization, which converts pixel values to float and scales the range from 0 to 255 to the range of 0 to 1, ensuring the data range is optimized for model convergence.

Visualization and Verification

Random examples from both the *Official* and *Wikipedia* classes were inspected to ensure the preprocessing steps were functioning correctly. An Official image that originally appeared as a high-resolution color document (for example, 3508 × 2480) was properly converted into the standardized 512 × 512 grayscale format used for model training. Likewise, a Wikipedia image with a different native size (such as 1752 × 1240) was also transformed into the same uniform 512 × 512 grayscale output. These checks confirm that the preprocessing pipeline reliably normalizes images regardless of their initial resolution or format.

Sample 1



DÉCLARATION DES ÉQUIPEMENTS

- N° Entreprise : □□□□□□□□

- N° Projet : □□□□□□□□

- Promoteur : ....

- Raison sociale : ....

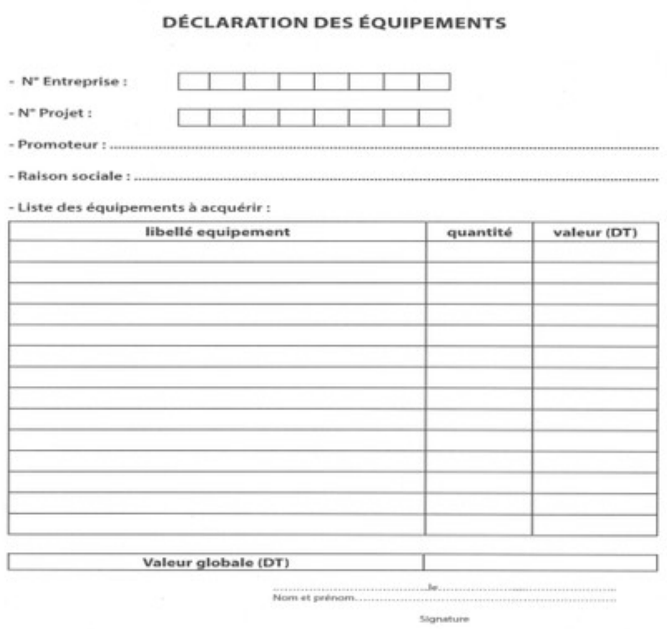
- Liste des équipements à acquérir :

libellé equipement	quantité	valeur (DT)

Valeur globale (DT)	

....le.......................................
Nom et prénom...

Signature



DÉCLARATION DES ÉQUIPEMENTS

- N° Entreprise : □□□□□□□□

- N° Projet : □□□□□□□□

- Promoteur : ....

- Raison sociale : ....

- Liste des équipements à acquérir :

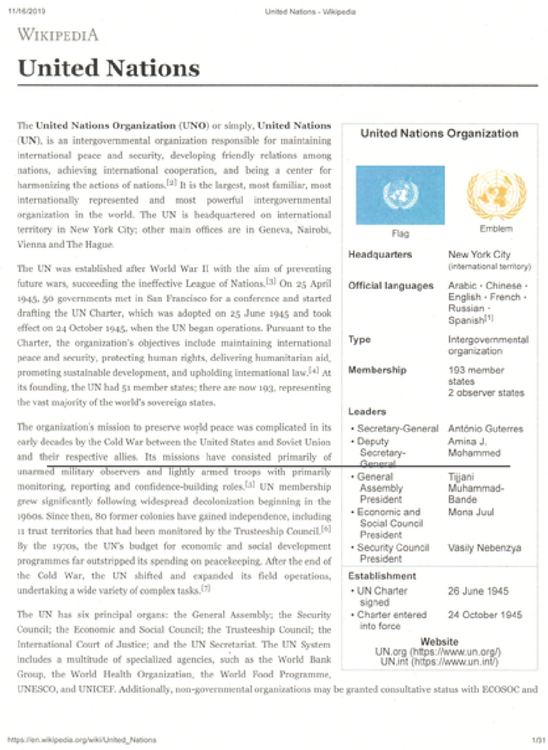
libellé equipement	quantité	valeur (DT)

Valeur globale (DT)	

....le.......................................
Nom et prénom...

Signature

Sample 2



11/16/2019 United Nations - Wikipedia

WIKIPEDIA

United Nations

The **United Nations Organization** (**UNO**) or simply, **United Nations** (**UN**), is an intergovernmental organization responsible for maintaining international peace and security, developing friendly relations among nations, achieving international cooperation, and being a center for harmonizing the actions of nations.[2] It is the largest, most familiar, most internationally represented and most powerful intergovernmental organization in the world. The UN is headquartered on international territory in New York City; other main offices are in Geneva, Nairobi, Vienna and The Hague.

The UN was established after World War II with the aim of preventing future wars, succeeding the ineffective League of Nations.[3] On 25 April 1945, 50 governments met in San Francisco for a conference and started drafting the UN Charter, which was adopted on 25 June 1945 and took effect on 24 October 1945, when the UN began operations. Pursuant to the Charter, the organization's objectives include maintaining international peace and security, protecting human rights, delivering humanitarian aid, promoting sustainable development, and upholding international law.[4] At its founding, the UN had 51 member states; there are now 193, representing the vast majority of the world's sovereign states.

The organization's mission to preserve world peace was complicated in its early decades by the Cold War between the United States and Soviet Union and their respective allies. Its missions have consisted primarily of unarmed military observers and lightly armed troops with primarily monitoring, reporting and confidence-building roles.[5] UN membership grew significantly following widespread decolonization beginning in the 1960s. Since then, 80 former colonies have gained independence, including 11 trust territories that had been monitored by the Trusteeship Council.[6] By the 1970s, the UN's budget for economic and social development programmes far outstripped its spending on peacekeeping. After the end of the Cold War, the UN shifted and expanded its field operations, undertaking a wide variety of complex tasks.[7]

The UN has six principal organs: the General Assembly; the Security Council; the Economic and Social Council; the Trusteeship Council; the International Court of Justice; and the UN Secretariat. The UN System includes a multitude of specialized agencies, such as the World Bank Group, the World Health Organization, the World Food Programme, UNESCO, and UNICEF. Additionally, non-governmental organizations may be granted consultative status with ECOSOC and

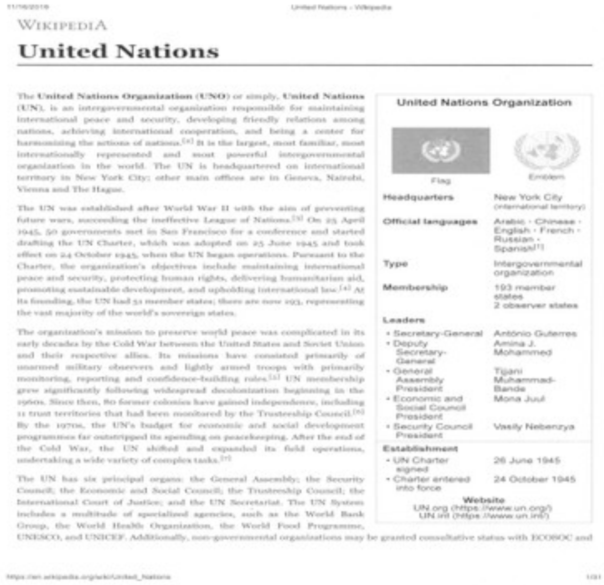
United Nations Organization

Flag Emblem

Headquarters	New York City (international territory)
Official languages	Arabic · Chinese · English · French · Russian · Spanish[1]
Type	Intergovernmental organization
Membership	193 member states 2 observer states
Leaders	
• Secretary-General	António Guterres
• Deputy Secretary-General	Amina J. Mohammed
• General Assembly President	Tijjani Muhammad-Bande
• Economic and Social Council President	Mona Juul
• Security Council President	Vasily Nebenzya
Establishment	
• UN Charter signed	26 June 1945
• Charter entered into force	24 October 1945

Website
UN.org (https://www.un.org/)
UN.int (https://www.un.int/)

https://en.wikipedia.org/wiki/United_Nations 1/31



11/16/2019 United Nations - Wikipedia

WIKIPEDIA

United Nations

The **United Nations Organization** (**UNO**) or simply, **United Nations** (**UN**), is an intergovernmental organization responsible for maintaining international peace and security, developing friendly relations among nations, achieving international cooperation, and being a center for harmonizing the actions of nations.[2] It is the largest, most familiar, most internationally represented and most powerful intergovernmental organization in the world. The UN is headquartered on international territory in New York City; other main offices are in Geneva, Nairobi, Vienna and The Hague.

The UN was established after World War II with the aim of preventing future wars, succeeding the ineffective League of Nations.[3] On 25 April 1945, 50 governments met in San Francisco for a conference and started drafting the UN Charter, which was adopted on 25 June 1945 and took effect on 24 October 1945, when the UN began operations. Pursuant to the Charter, the organization's objectives include maintaining international peace and security, protecting human rights, delivering humanitarian aid, promoting sustainable development, and upholding international law.[4] At its founding, the UN had 51 member states; there are now 193, representing the vast majority of the world's sovereign states.

The organization's mission to preserve world peace was complicated in its early decades by the Cold War between the United States and Soviet Union and their respective allies. Its missions have consisted primarily of unarmed military observers and lightly armed troops with primarily monitoring, reporting and confidence-building roles.[5] UN membership grew significantly following widespread decolonization beginning in the 1960s. Since then, 80 former colonies have gained independence, including 11 trust territories that had been monitored by the Trusteeship Council.[6] By the 1970s, the UN's budget for economic and social development programmes far outstripped its spending on peacekeeping. After the end of the Cold War, the UN shifted and expanded its field operations, undertaking a wide variety of complex tasks.[7]

The UN has six principal organs: the General Assembly; the Security Council; the Economic and Social Council; the Trusteeship Council; the International Court of Justice; and the UN Secretariat. The UN System includes a multitude of specialized agencies, such as the World Bank Group, the World Health Organization, the World Food Programme, UNESCO, and UNICEF. Additionally, non-governmental organizations may be granted consultative status with ECOSOC and

United Nations Organization

Flag Emblem

Headquarters	New York City (international territory)
Official languages	Arabic · Chinese · English · French · Russian · Spanish[1]
Type	Intergovernmental organization
Membership	193 member states 2 observer states
Leaders	
• Secretary-General	António Guterres
• Deputy Secretary-General	Amina J. Mohammed
• General Assembly President	Tijjani Muhammad-Bande
• Economic and Social Council President	Mona Juul
• Security Council President	Vasily Nebenzya
Establishment	
• UN Charter signed	26 June 1945
• Charter entered into force	24 October 1945

Website
UN.org (https://www.un.org/)
UN.int (https://www.un.int/)

https://en.wikipedia.org/wiki/United_Nations 1/31

Conclusion

Milestone 1 has been successfully wrapped up. The entire dataset has been examined, cataloged, and processed, resulting in a fully standardized collection of NumPy arrays with the shape (N, 512, 512, 1). All images are now uniform in resolution, converted to grayscale, and properly normalized. With this foundation in place, the project is ready to move forward to Milestone 2, which will focus on feature engineering and building initial baseline models.