

# Supplementary Material for

## SmolDocling: An ultra-compact vision-language model for end-to-end multi-modal document conversion

Ahmed Nassar Matteo Omenetti Maksym Lysak Nikolaos Livathinos Christoph Auer Lucas Morin  
 Rafael Teixeira de Lima Yusik Kim A. Said Gurbuz Michele Dolfi Peter W. J. Staar  
 {ahn, mly, nli, cau, dol, taa}@zurich.ibm.com, {Matteo.Omenetti1}@ibm.com  
 IBM Research

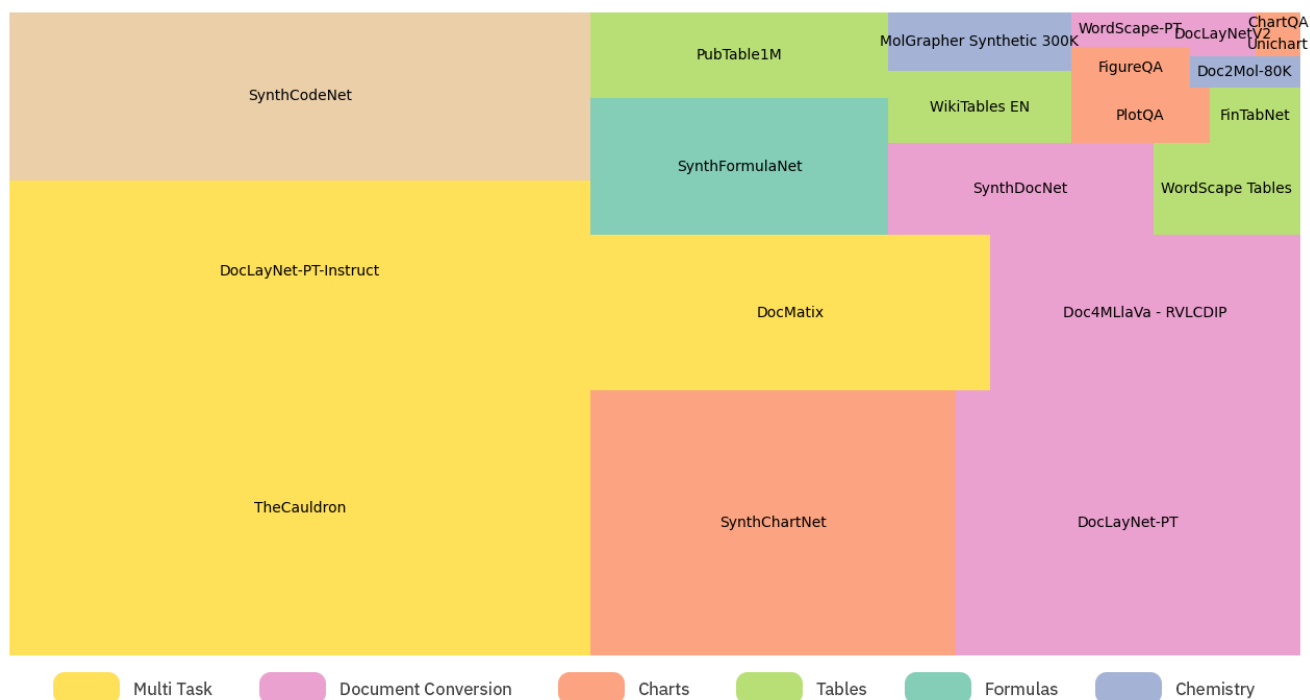


Figure 1. **Training datasets.** A treemap visualization of dataset sizes contributing to the training of SmolDocling, categorized by type. Each rectangle represents a dataset, with its area proportional to the total dataset size. Each color indicates a dataset type. “Multi-Task” refers to datasets that include multiple different tasks.

## 1. Additional Datasets Details

### 1.1. SynthChartNet

SynthChartNet includes four distinct types of charts: line, pie, bar, and stacked bar, generated using three visualization libraries: Matplotlib [? ], Seaborn [? ], and Pyecharts [? ]. Each data set was rendered through all three libraries, producing a final dataset of 2.5 million visually diverse charts. The rendering process was designed to leverage the full range of features available in these libraries, ensuring visual variety while maintaining the plausibility of the data presented. Figure 2 depicts 10 samples from SynthChartNet.

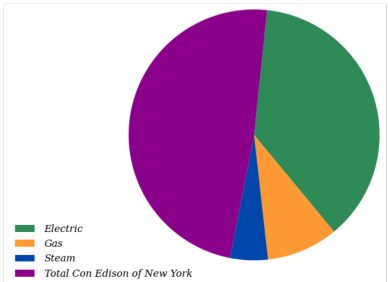
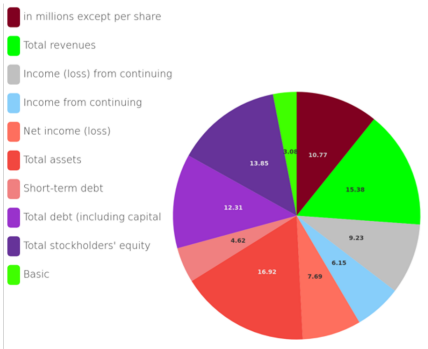
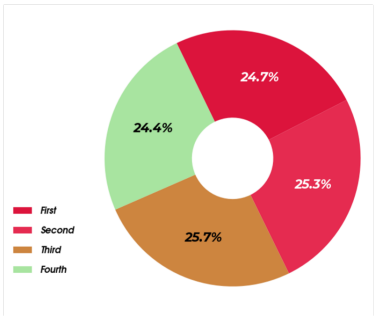
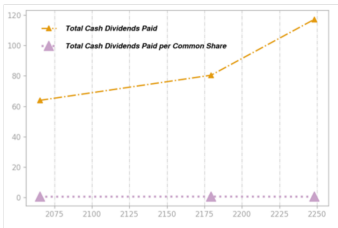
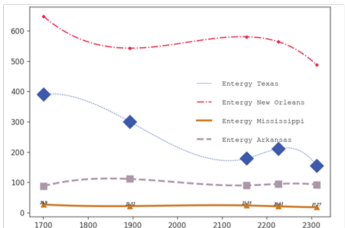
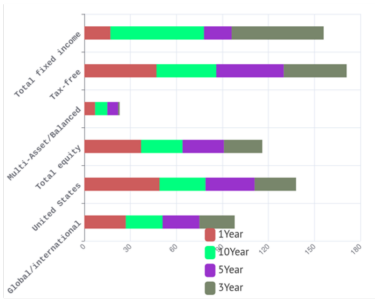
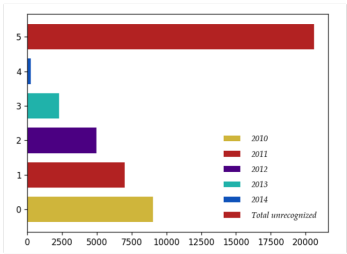
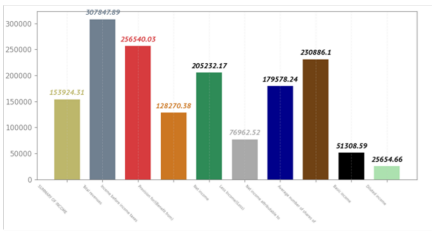


Figure 2. Samples from SynthChartNet.

```

1. #pragma once
2.
3. #include <sdm/ utils/ config. hpp>
4. #include <sdm/ utils/ struct/ vector. hpp>
5. #include <sdm/ world/ solvable_ by_ dp. hpp>
6. #include <sdm/ utils/ value_ function/ initializer. hpp>
7.
8. namespace sdm
9. {
10.     namespace initializer
11.     {
12.         class registry
13.         {
14.         protected:
15.             typedef std::map<std::string, std::shared_ptr<
16.                 Initializer> (*) (std::shared_ptr<SolvableByDP>
17.                     world, Config config)> map_type;
18.             static map_type container;
19.
20.         public:
21.             static std::vector<std::string> available();
22.             static std::shared_ptr<Initializer> make(std::
23.                 string name, std::shared_ptr<SolvableByDP>
24.                     world, Config config = {});
25.         };
26.
27.         template <class TInitializer>
28.         std::shared_ptr<Initializer> createInstance(std::
29.             shared_ptr<SolvableByDP> world, Config config)
30.         {
31.             return std::make_shared<TInitializer>(world,
32.                 config);
33.         }
34.     }
35. }

```

```

package br.com.biot.integracaopagarmeapi.modulos.jwt.utis

import br.com.biot.integracaopagarmeapi.config.exception.ValidacaoException
import io.jsonwebtoken.Claims
import lombok.extern.slf4j.Slf4j

@Slf4j
public class JwtCampoUtil {

    public static String getCampoId(Claims claims) {
        try {
            return (String) claims.get("id");
        } catch (Exception ex) {
            log.error("Erro ao tentar recuperar ID do JWT: {}", claims.toString());
            throw new ValidacaoException("Erro ao tentar recuperar o campo ID do Token.");
        }
    }

    public static String getCampo(String campo, Claims claims) {
        try {
            return (String) claims.get(campo);
        } catch (Exception ex) {
            log.error("Erro ao tentar recuperar o campo {} do JWT: {}", campo, claims.toString());
            throw new ValidacaoException(
                String.format("Erro ao tentar recuperar o campo %s do Token.", campo)
            );
        }
    }
}

```

```

name: Stage the app

on:
  pull_request:
    types: [labeled]

env:
  DOCKER_IMAGE_NAME: yelimell-azure-rtt
  IMAGE_REGISTRY_URL: docker.pkg.github.com
  #####
  *** USER PROVIDED VALUES ARE REQUIRED BELOW ***
  #####
  *** REPLACE USERNAME WITH ON VIDEOSNAME ***
  AZURE_REGISTRY_NAME: yelimell-rtt-app
  #####

jobs:
  build:
    if: contains(github.event.pull_request.labels.*.name, 'stage')
    runs-on: ubuntu-latest

    steps:
      - uses: actions/checkout@v1
      - name: npm install and build webpack
        run: |
          npm install
          npm run build
      - uses: actions/upload-artifact@main
        with:
          name: webpack artifacts
          path: public/

  Build-Docker-Image:
    runs-on: ubuntu-latest
    needs: build
    name: Build image and store in GitHub Packages
    steps:
      - name: Checkout
        uses: actions/checkout@v1
      - name: Download built artifact
        uses: actions/download-artifact@main
        with:
          name: webpack artifacts
          path: public
      - name: create image and store in Packages
        uses: mattdevia031/actions/docker-gpr@1.3.0
        with:
          repository: ${{ secrets.GITHUB_TOKEN }}
          image-name: ${{ env.DOCKER_IMAGE_NAME }}

  Deploy-to-Azure:
    runs-on: ubuntu-latest
    needs: Build-Docker-Image
    name: Deploy app container to Azure
    steps:
      - name: "Login via Azure CLI"
        uses: azure/login@v1
        with:
          creds: ${{ secrets.AZURE_CREDENTIALS }}
      - uses: azure/docker-login@v1
        with:
          login-server: ${{ env.IMAGE_REGISTRY_URL }}
          username: ${{ github.actor }}

```

```

1 def main(filename, days):
2     # array of buckets 0-8, each representing the number of fish
3     # with that counter value
4     fish_by_counter = [0 for i in range(9)]
5     for counter in read_counters(filename):
6         fish_by_counter[counter] += 1
7
8     for i in range(days):
9         fish_by_counter = spawn(fish_by_counter)
10
11     print(f"After {days} days: {sum(fish_by_counter)} fish")
12
13 def spawn(fish_by_counter):
14     # rotate left
15     res = fish_by_counter[1:] + fish_by_counter[:1]
16
17     # fish that were in bucket 0 are now new fish in bucket 8.
18     # restore the original fish into bucket 6.
19     res[6] += res[8]
20     return res
21
22 def read_counters(filename):
23     with open(filename) as f:
24         return [int(counter) for counter in f.readline().strip().
25             split(',')]
26
27 if __name__ == '__main__':
28     main('input.txt', 80)
29     main('input.txt', 256)

```

Figure 3. Samples from SynthCodeNet.

## 1.2. SynthCodeNet

To render the code snippets of SynthCodeNet, we employed two rendering libraries: LaTeX and Pygments [? ]. LaTeX, in conjunction with its listings package, enabled the generation of code renderings similar to those commonly seen in academic and technical publications, while Pygments facilitated renderings with aesthetics resembling those of IDEs. The rendering process utilized the complete set of features provided by these libraries to enhance visual diversity while preserving the plausibility of the presented data. This included randomizing various stylistic parameters such as font, font size, style, weight, line spacing, tab size, background color, line numbering (including randomizing the starting line number), and programming language-specific syntax highlighting. Figure 3 shows 4 samples from SynthCodeNet.

$$\begin{aligned}
\sum_{k \in f \cap K_m} e^{-i\varphi_k} m_{2Q_k}(e^{i\varphi_k} f) \nabla \chi_k &= \sum_{k \in f \cap K_m} \left( e^{-i\varphi_k} m_{2Q_k}(e^{i\varphi_k} f) - e^{-i\tilde{\varphi}_m} m_{2Q_k}(e^{i\tilde{\varphi}_m} f) \right) \nabla \chi_k \\
&+ \sum_{k \in f \cap K_m} e^{-i\tilde{\varphi}_m} \left( m_{2Q_k}(e^{i\tilde{\varphi}_m} f) - m_{\tilde{Q}_m}(e^{i\tilde{\varphi}_m} f) \right) \nabla \chi_k \\
&+ \sum_{k \in f \cap K_m} e^{-i\tilde{\varphi}_m} m_{\tilde{Q}_m}(e^{i\tilde{\varphi}_m} f) \nabla \chi_k \quad \text{Y} \\
&= I + II + III.
\end{aligned}$$

$$Rd \leq \alpha \ln \left( \frac{(A_y + \lambda_y)^{A_y + \lambda_y}}{(A_z + \lambda_z)^{A_z + \lambda_z}} \right) + (1 - \alpha) \ln \left( \frac{\lambda_y^{\lambda_y}}{\lambda_z^{\lambda_z}} \right) - \ln \left( \frac{(A_y \alpha + \lambda_y)^{A_y \alpha + \lambda_y}}{(A_z \alpha + \lambda_z)^{A_z \alpha + \lambda_z}} \right) \quad 4.30$$

$$R \leq \alpha \ln \left( (A_y + \lambda_y)^{A_y + \lambda_y} \right) + (1 - \alpha) \ln \left( \lambda_y^{\lambda_y} \right) - \ln \left( (A_y \alpha + \lambda_y)^{A_y \alpha + \lambda_y} \right) \quad 4.31$$

$$d \leq 1 \quad 4.32$$

$$u_{ij}^{(0)} = \begin{cases} \frac{1}{(1+q)^{j-1}}, & (i=1), \\ \frac{q}{(1+q)^{j-i+1}}, & (2 \leq i \leq j), \\ 0, & (j+1 \leq i \leq N). \end{cases} \quad (\text{L})$$

$$d\Phi=\left(\frac{2\pi}{k}\right)^2s_z^2\bar{a}(-\rho s)\bar{a}^*(-\rho s)\,d\Omega.$$

$$\begin{aligned}
\mathcal{L}_1 &= \left\{ \max(\frac{\bar{a}}{\bar{b}}\gamma, \frac{\gamma(\bar{a}+b) - \sqrt{\gamma^2(\bar{a}+b)^2 - 4a\bar{b}}}{2\bar{b}}) \leq \beta \leq \frac{\gamma(\bar{a}+b) + \sqrt{\gamma^2(\bar{a}+b)^2 - 4a\bar{b}}}{2\bar{b}} \right\} \\
\mathcal{L}_2 &= \left\{ \frac{(a+\bar{b}) + \sqrt{(a+\bar{b})^2 - 4\bar{a}b\gamma^2}}{2b\gamma} \leq \beta \leq \frac{\bar{a}}{\bar{b}}\gamma \right\} \\
\mathcal{L}_3 &= \left\{ \beta \leq \min(\frac{\bar{a}}{\bar{b}}\gamma, \frac{(a+\bar{b}) - \sqrt{(a+\bar{b})^2 - 4\bar{a}b\gamma^2}}{2b\gamma}) \right\} \\
\mathcal{L}_4 &= \left\{ \beta \leq \min(\frac{b\gamma}{a}, \frac{\gamma(\bar{a}+b) - \sqrt{\gamma^2(\bar{a}+b)^2 - 4a\bar{b}}}{2a}) \right\} \\
\mathcal{L}_5 &= \left\{ \frac{\gamma(\bar{a}+b) + \sqrt{\gamma^2(\bar{a}+b)^2 - 4a\bar{b}}}{2a} \leq \beta \leq \frac{b\gamma}{a} \right\} \\
\mathcal{L}_6 &= \left\{ \max(\frac{b\gamma}{a}, \frac{(a+\bar{b}) - \sqrt{(a+\bar{b})^2 - 4\bar{a}b\gamma^2}}{2\bar{a}\gamma}) \leq \beta \leq \frac{(a+\bar{b}) + \sqrt{(a+\bar{b})^2 - 4\bar{a}b\gamma^2}}{2\bar{a}\gamma} \right\}, \quad (\text{x})
\end{aligned}$$

$$\begin{aligned}
\int_{S^{n-1}} v(s) \, d\mu^\sigma &\geq \frac{(n-2)\omega_{n-1}}{2} m^{U^\kappa b} \\
&+ \int_s^\infty \left( 1 - \frac{\cosh s \, u_0(r)}{\cosh r \, u_0(s)} \right) \cosh r \sinh^{n-1} r \left( \int_{S^{n-1}} f(u(r, \theta)) \, d\mu^\sigma \right) dr
\end{aligned} \quad (21.64)$$

Figure 4. Samples from SynthFormulaNet.



### 1.3. SynthFormulaNet

The extracted formulas of SynthFormulaNet underwent a rigorous normalization procedure to ensure the model was trained only on correct and standardized code. The normalization process consisted of the following major steps:

1. **Token Filtering and Removal:** Unnecessary or redundant LaTeX tokens were removed or replaced based on predefined policies. This included:
  - Eliminating unwanted tokens.
  - Replacing tokens with equivalent ones for consistency.
2. **Structural Normalization:** The equation structure was standardized by:
  - Normalizing spacing between tokens.
  - Standardizing left and right delimiters.
  - Ensuring consistent use of tokens that require one or two braces.
3. **Token Simplification:** Redundant tokens and patterns were simplified, including:
  - Collapsing consecutive redundant tokens (e.g., repeated primes, dots, or excessive spacing).
  - Converting special constructs such as `\Big` into `\left` and `\right` for consistency.

The extracted equations were subsequently rendered using LaTeX at 120 dpi. Several aspects of the rendering procedure have been randomized to ensure the visual diversity of the dataset. Each formula was rendered using a randomly selected font from a set of 85 possible options, with variations in font size, style, weight, and line spacing. Additionally, with a certain probability, each formula was enclosed within an "align" environment, allowing for the inclusion of a randomly generated equation number on either the left or right side of the equation. The equation number itself was chosen at random from a diverse set of possibilities, including small and large numbers, integers, floating-point values, or letters. Furthermore, it could be enclosed in single or double parentheses, brackets, or curly braces and was sometimes preceded by text labels such as "Eqn.", "Eqn", "Eq.", or "Equation". [Figure 4](#) depicts 6 samples from SynthFormulaNet.

### 1.4. DocTags vocabulary

Here we provide a reference to a complete vocabulary of *DocTags* for *SmolDocling*, and their meaning.

We represent tags in an XML-like style of notation. Some tags have an opening and closing part, and wrap around textual content (or other tags) for example: `<text>hello world</text>`. Other tags are standalone and do not have a closing tag; they mark a special instruction, for example: `<page.break>`

The complete *DocTags* snippet can represent a page or multiple pages, wrapped into `<doctag>...</doctag>`. When content represents multiple pages, we use `<page.break>` tag as a separator.

Within the scope of `<doctag>` we place high-level, tags that wrap around the textual content of certain document blocks and identify the type of the block. These tags are: `<text>`, `<caption>`, `<footnote>`, `<formula>`, `<title>`, `<page.footer>`, `<page.header>`, `<picture>`, `<section.header>`, `<document.index>`, `<code>`, `<otsl>`, `<list.item>`, `<ordered.list>`, `<unordered.list>`

Each element can nest additional standalone location tags that encode its position on the page as a bounding box, represented in DocTags as `<loc_x1><loc_y1><loc_x2><loc_y2>`. Every x,y pair of coordinates belongs to a fixed grid. In our case we used integer values in the range from 0 to 500 which proportionally maps to width and height of the page.

`<otsl>` element contains table representation with following OTSL tags, that are standalone and used to describe structure of a table: `<fcel>`, `<ecel>`, `<lcel>`, `<ucel>`, `<xcel>`, `<n1>`. These tags follow rules as described in [? ], we additionally differentiate OTSL notion of a cell into `<fcel>` - full cell, or cell which contain content, and `<ecel>` - empty cell. We also augment tabular structure information with extra tags `<ched>`, `<rhed>`, `<srow>` that replace `<fcel>` to describe cells which belong to column and row headers of the table, as well as table section accordingly. We mark such headers when ground truth about table headers from the underlying dataset is available.

`<list.item>` elements are placed within `<ordered.list>` or `<unordered.list>` and define whether it's enumerated (ordered) list or not.

`<picture>` and `<otsl>` elements by themselves can encapsulate `<caption>` tag if appropriate picture or table has it's own caption. This way we can connect extra descriptive information to illustrations and tables in the document.

`<code>` elements contain pieces of code as a content, and as such, respect tabulation and line breaks. Code elements also contains special standalone classification tag - `<_programming-language->` where *programming-language* is an appropriate programming language name. Supported values (57 in total) for programming language names are: *Ada*, *Awk*, *Bash*, *bc*, *C*, *C#*, *C++*, *CMake*, *COBOL*, *CSS*, *Ceylon*, *Clojure*, *Crystal*, *Cuda*, *Cython*, *D*, *Dart*, *dc*, *Dockerfile*, *Elixir*, *Erlang*, *FORTRAN*, *Forth*, *Go*, *HTML*, *Haskell*, *Haxe*, *Java*, *JavaScript*, *Julia*, *Kotlin*, *Lisp*, *Lua*, *Matlab*, *MoonScript*, *Nim*, *OCaml*,

ObjectiveC, Octave, PHP, Pascal, Perl, Prolog, Python, Racket, Ruby, Rust, SML, SQL, Scala, Scheme, Swift, TypeScript, unknown, VisualBasic, XML, YAML

Within `<picture>` elements we include picture classification with extra standalone tags `<image_class>`. We classify images from our datasets into following categories: *natural\_image*, *pie\_chart*, *bar\_chart*, *line\_chart*, *flow\_chart*, *scatter\_chart*, *heatmap*, *remote\_sensing*, *chemistry\_molecular\_structure*, *chemistry\_markush\_structure*, *icon*, *logo*, *signature*, *stamp*, *qr\_code*, *bar\_code*, *screenshot*, *map*, *stratigraphic\_chart*, *cad\_drawing*, *electrical\_diagram*

2. Additional Details about the SmolDocling

| Instruction               | Description  |
|---------------------------|--|
| Full conversion           | Convert this page to docling.  |
| Chart                     | Convert chart to table (e.g., <code>&lt;chart&gt;</code> ).  |
| Formula                   | Convert formula to LaTeX (e.g., <code>&lt;formula&gt;</code> ).  |
| Code                      | Convert code to text (e.g., <code>&lt;code&gt;</code> ).   |
| Table                     | Convert table to OTSL (e.g., <code>&lt;otsl&gt;</code> ).  |
| No-Code Actions/Pipelines | OCR the text in a specific location: <code>&lt;loc.155&gt;&lt;loc.233&gt;&lt;loc.206&gt;&lt;loc.237&gt;</code> |
|                           | Identify element at: <code>&lt;loc.247&gt;&lt;loc.482&gt;&lt;loc.252&gt;&lt;loc.486&gt;</code>                 |
|                           | Find all ‘text’ elements on the page, retrieve all section headers.  |
|                           | Detect footer elements on the page.  |

Table 1. Supported Conversion Instructions.

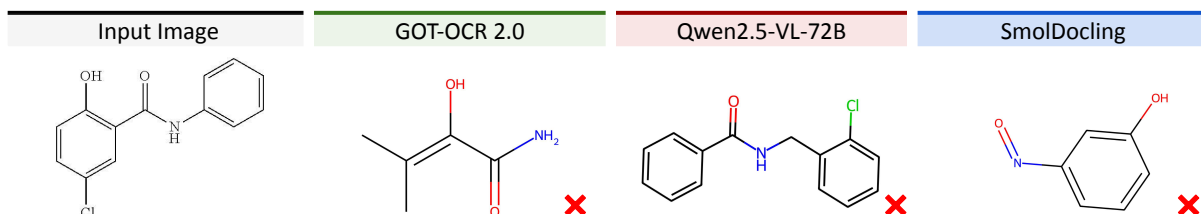


Figure 5. **Qualitative comparison for molecule image recognition.** Predictions are shown for GOT-OCR 2.0, Qwen2.5-VL-72B and SmolDocling on a simple molecule image.

### 3. Additional Qualitative Analysis

#### 3.1. Molecule Image Recognition Experiment

Extracting molecule images from documents has potential to accelerate research in chemistry and materials discovery [? ]. Some recent document understanding models, such as GOT-OCR 2.0 [? ] and Qwen2.5-VL [? ], claim to perform molecule image recognition in documents. Following this direction, we experimented with training SmolDocling for molecule image recognition on full document pages and cropped images. In order to predict molecule structures (graphs) with SmolDocling, we use the molecule string identifier named SMILES [? ].

**Datasets.** For training on cropped images, we use the MolGrapher-Synthetic-300K dataset [? ]. The dataset contains 300K synthetic samples. It is created using real chemical-structures retrieved from the database PubChem [? ] which are then synthetically drawn using the library RDKit [? ]. For training of full pages, we use Doc2Mol-80K, a subset of the PatCID dataset [? ]. Our subset contains 80 000 real patent documents published in the United-States between 2022 and 2024.

**Evaluation.** Figure 5 shows an example of predicted molecules for a simple molecule image. While some models capture high-level molecular features, they consistently fail to reconstruct its detailed structure. Notably, this molecule is trivial for specialized models such as MolGrapher [? ], MolScribe [? ] or DECIMER [? ], which accurately recognize its structure. Reducing the gap between general document understanding models and specialized models may require incorporating a specialized vocabulary for SMILES tokens, explicitly encoding atoms and bonds, and extending the SMILES sequence with localization tokens for each atom.

#### 3.2. Layout analysis samples

To illustrate and compare the visual grounding capabilities of SmolDocling, a set of sample predictions with SmolDocling and Qwen2.5-VL from DocLayNet [? ] is shown in Table 2. Note that the element location results are independent from the correct reproduction of document content and structure.

| # | Ground-Truth   | SmolDocling   | Qwen2.5-VL   | ID                                |
|---|--|---|--|-----------------------------------|
| 1 | <div data-bbox="224 262 587 814" data-label="Image"> </div>  | <div data-bbox="662 262 1026 814" data-label="Image"> </div>  | <div data-bbox="1101 262 1464 814" data-label="Image"> </div>  | 0b49b1c3d5217411e7c9b9d2960d5f61  |
| 2 | <div data-bbox="224 877 587 1348" data-label="Image"> </div> | <div data-bbox="662 877 1026 1348" data-label="Image"> </div> | <div data-bbox="1101 877 1464 1348" data-label="Image"> </div> | 0aafe8d4b66ac7ab81e9456327fb2db01 |

Continued on next page

| # | Ground-Truth  | SmolDocling  | Qwen2.5-VL   | ID                               |
|---|---|--|--|----------------------------------|
| 3 | <p><b>11</b> Workers engaged in essential or emergency works—</p> <p>(a) related to water supplies and sewerage services; and</p> <p>(b) carried out by, for, or on behalf of a water undertaker, sewerage undertaker, water supply licensee, sewerage licensee or local authority.</p> <p><b>12</b> They have travelled to the United Kingdom in the course of their work.</p> <p><b>13</b> For the purposes of sub-paragraph (1)—</p> <p>(a) “essential or emergency works” includes—</p> <p>(i) inspections, maintenance, repairs, and asset replacement activities;</p> <p>(ii) monitoring, sampling and analysis of water supplies under the Private Water Supplies (England) Regulations 2016(a), the Water Supply (Water Quality) Regulations 2016(b), the Private Water Supplies (Wales) Regulations 2017(a), or the Water Supply (Water Quality) Regulations 2016(b);</p> <p>(iii) “sewerage licensee” means the holder of a sewerage licence under section 178A of the Water Industry Act 1991(c);</p> <p>(iv) “sewerage services” has the meaning given in section 219(1) of the Water Industry Act 1991(c);</p> <p>(v) “water supply licensee” has the meaning given in sections 17A(7) and 219(1) of the Water Industry Act 1991(c);</p> <p><b>14</b> Workers engaged in essential or emergency works relating to flood and coastal erosion risk management on behalf of—</p> <p>(a) the Environment Agency; or</p> <p>(b) a local flood authority in England;</p> <p><b>15</b> For the purposes of sub-paragraph (1)—</p> <p>(a) “flood” and “coastal erosion” have the meanings given in section 1 of the Flood and Water Management Act 2010(d);</p> <p>(b) “local flood authority” has the meaning given in section 6(7) of that Act;</p> <p>(c) “risk management” has the meaning given in section 7 of that Act;</p> <p><b>16</b> Workers engaged in essential or emergency works—</p> <p>(a) related to—</p> <p>(i) generating stations;</p> <p>(ii) electricity interconnectors;</p> <p>(iii) a district heat network as defined in regulation 2 of the Heat Network (Metering and Billing) Regulations 2014(g);</p> <p>(iv) central heating as defined in regulation 2 of the Heat Network (Metering and Billing) Regulations 2014;</p> <p>(v) industrial ballast cleaning and track re-laying systems on a network; or</p> <p>(vi) commissioning, maintenance and repair of industrial machinery for use on a network; or</p> <p><b>17</b> There are no relevant preceding provisions.</p>   | <p><b>11</b> Workers engaged in essential or emergency works—</p> <p>(a) related to water supplies and sewerage services; and</p> <p>(b) carried out by, for, or on behalf of a water undertaker, sewerage undertaker, water supply licensee, sewerage licensee or local authority.</p> <p><b>12</b> They have travelled to the United Kingdom in the course of their work.</p> <p><b>13</b> For the purposes of sub-paragraph (1)—</p> <p>(a) “essential or emergency works” includes—</p> <p>(i) inspections, maintenance, repairs, and asset replacement activities;</p> <p>(ii) monitoring, sampling and analysis of water supplies under the Private Water Supplies (England) Regulations 2016(a), the Water Supply (Water Quality) Regulations 2016(b), the Private Water Supplies (Wales) Regulations 2017(a), or the Water Supply (Water Quality) Regulations 2016(b);</p> <p>(iii) “sewerage licensee” means the holder of a sewerage licence under section 178A of the Water Industry Act 1991(c);</p> <p>(iv) “sewerage services” has the meaning given in section 219(1) of the Water Industry Act 1991(c);</p> <p>(v) “water supply licensee” has the meaning given in sections 17A(7) and 219(1) of the Water Industry Act 1991(c);</p> <p><b>14</b> Workers engaged in essential or emergency works relating to flood and coastal erosion risk management on behalf of—</p> <p>(a) the Environment Agency; or</p> <p>(b) a local flood authority in England;</p> <p><b>15</b> For the purposes of sub-paragraph (1)—</p> <p>(a) “flood” and “coastal erosion” have the meanings given in section 1 of the Flood and Water Management Act 2010(d);</p> <p>(b) “local flood authority” has the meaning given in section 6(7) of that Act;</p> <p>(c) “risk management” has the meaning given in section 7 of that Act;</p> <p><b>16</b> Workers engaged in essential or emergency works—</p> <p>(a) related to—</p> <p>(i) generating stations;</p> <p>(ii) electricity interconnectors;</p> <p>(iii) a district heat network as defined in regulation 2 of the Heat Network (Metering and Billing) Regulations 2014(g);</p> <p>(iv) central heating as defined in regulation 2 of the Heat Network (Metering and Billing) Regulations 2014;</p> <p>(v) industrial ballast cleaning and track re-laying systems on a network; or</p> <p>(vi) commissioning, maintenance and repair of industrial machinery for use on a network; or</p> <p><b>17</b> There are no relevant preceding provisions.</p>  | <p><b>21.—(1)</b> Workers engaged in essential or emergency works—</p> <p>(a) related to water supplies and sewerage services; and</p> <p>(b) carried out by, for, or on behalf of a water undertaker, sewerage undertaker, water supply licensee, sewerage licensee or local authority.</p> <p><b>22</b> They have travelled to the United Kingdom in the course of their work.</p> <p><b>23</b> For the purposes of sub-paragraph (1)—</p> <p>(a) “essential or emergency works” includes—</p> <p>(i) inspections, maintenance, repairs, and asset replacement activities;</p> <p>(ii) monitoring, sampling and analysis of water supplies under the Private Water Supplies (England) Regulations 2016(a), the Water Supply (Water Quality) Regulations 2016(b), the Private Water Supplies (Wales) Regulations 2017(a), or the Water Supply (Water Quality) Regulations 2016(b);</p> <p>(iii) “sewerage licensee” means the holder of a sewerage licence under section 178A of the Water Industry Act 1991(c);</p> <p>(iv) “sewerage services” has the meaning given in section 219(1) of the Water Industry Act 1991(c);</p> <p>(v) “water supply licensee” has the meaning given in sections 17A(7) and 219(1) of the Water Industry Act 1991(c);</p> <p><b>24</b> Workers engaged in essential or emergency works relating to flood and coastal erosion risk management on behalf of—</p> <p>(a) the Environment Agency; or</p> <p>(b) a local flood authority in England;</p> <p><b>25</b> For the purposes of sub-paragraph (1)—</p> <p>(a) “flood” and “coastal erosion” have the meanings given in section 1 of the Flood and Water Management Act 2010(d);</p> <p>(b) “local flood authority” has the meaning given in section 6(7) of that Act;</p> <p>(c) “risk management” has the meaning given in section 7 of that Act;</p> <p><b>26</b> Workers engaged in essential or emergency works—</p> <p>(a) related to—</p> <p>(i) generating stations;</p> <p>(ii) electricity interconnectors;</p> <p>(iii) a district heat network as defined in regulation 2 of the Heat Network (Metering and Billing) Regulations 2014(g);</p> <p>(iv) central heating as defined in regulation 2 of the Heat Network (Metering and Billing) Regulations 2014;</p> <p>(v) industrial ballast cleaning and track re-laying systems on a network; or</p> <p>(vi) the commissioning, maintenance and repair of industrial machinery for use on a network; or</p> <p><b>27</b> There are no relevant preceding provisions.</p>   | 1bfr94Bce34e976dc3aabc5b8b43f355 |
| 4 | <p><b>18</b> Equations of the saddle point are the following</p> $C - \beta^2 + q \log(a_1^2 + a_2^2) + \int D^2 \phi(x) \log  x  - \beta^2 = 0 \quad (11)$ <p><b>19</b> <math>C</math> is a Lagrange multiplier enforcing the constraint (9). These equations are valid on the support of <math>\rho</math>. We can prove that these equations lead to a singular support for <math>\rho</math>. We do this by contradiction.</p> <p><b>20</b> Assume that a smooth locus for the support of <math>\rho</math> is allowed. In that case, we can differentiate the equation 11 six times, with the Laplacian in 6 dimensions to the cubed power. We find then that</p> $(\nabla^2)^3 q \log(a_1^2 + a_2^2) + \int D^2 \phi(x) \nabla^2 \log  x  - \beta^2 = 0 \quad (12)$ <p><b>21</b> The first two terms obviously are cancelled. Notice that</p> $\frac{\nabla^2(a_1^2 + a_2^2)}{\log(a_1^2 + a_2^2)} \approx \log(x^2) + \log(x^2) \quad (13)$ <p><b>22</b> That away from <math>x^2 = 0</math> we have that</p> $\nabla^2 \log(a_1^2 + a_2^2) = \nabla^2 (\log(x^2) + \log(x^2)) = 0 \quad (14)$ <p><b>23</b> The function is harmonic: a sum of a holomorphic plus an antiholomorphic piece. Also, it is easy to show that <math>(\nabla^2)^3 \log(x^2 - \delta) \propto \delta^2(x^2 - \delta)</math>. With these, we would find that <math>\beta = 0</math>. This argument contradicts that <math>\beta \neq 0</math>, so the assumption that <math>\rho</math> is smooth does not work. This argument has been given before in [5, 25], and we are presenting it here for completeness.</p> <p><b>24</b> That this indicator is that the density distribution of particles is singular. Since for <math>q = 0</math> the original probability distribution is spherically symmetric, it is natural to assume that the associated distribution is spherically symmetric. The radius was computed in [18]. This has also been checked numerically in [6, 7]. A generalisation to all Calde-Yau cone conformal field theory singularities for this result can be found in [25], where it was seen that this procedure matches various other supergravity calculations in the absence of symmetry.</p> <p><b>25</b> For <math>q &gt; 0</math> but very small, we expect that the distribution does not change much, except in the region <math>x^2 \approx 0</math>. Thus we expect that the distribution gives a five dimensional manifold of <math>\mathbb{R}^5</math>, possibly with a boundary, and that the region near <math>x^2 = 0</math> has no density of eigenvalues present. After all, the effective potential repels all particles from the locus <math>x^2 = 0</math>. This</p> | <p><b>18</b> Equations of the saddle point are the following</p> $C - \beta^2 + q \log(a_1^2 + a_2^2) + \int D^2 \phi(x) \log  x  - \beta^2 = 0 \quad (11)$ <p><b>19</b> <math>C</math> is a Lagrange multiplier enforcing the constraint (9). These equations are valid on the support of <math>\rho</math>. We can prove that these equations lead to a singular support for <math>\rho</math>. We do this by contradiction.</p> <p><b>20</b> Assume that a smooth locus for the support of <math>\rho</math> is allowed. In that case, we can differentiate the equation 11 six times, with the Laplacian in 6 dimensions to the cubed power. We find then that</p> $(\nabla^2)^3 q \log(a_1^2 + a_2^2) + \int D^2 \phi(x) \nabla^2 \log  x  - \beta^2 = 0 \quad (12)$ <p><b>21</b> The first two terms obviously are cancelled. Notice that</p> $\frac{\nabla^2(a_1^2 + a_2^2)}{\log(a_1^2 + a_2^2)} \approx \log(x^2) + \log(x^2) \quad (13)$ <p><b>22</b> That away from <math>x^2 = 0</math> we have that</p> $\nabla^2 \log(a_1^2 + a_2^2) = \nabla^2 (\log(x^2) + \log(x^2)) = 0 \quad (14)$ <p><b>23</b> The function is harmonic: a sum of a holomorphic plus an antiholomorphic piece. Also, it is easy to show that <math>(\nabla^2)^3 \log(x^2 - \delta) \propto \delta^2(x^2 - \delta)</math>. With these, we would find that <math>\beta = 0</math>. This argument contradicts that <math>\beta \neq 0</math>, so the assumption that <math>\rho</math> is smooth does not work. This argument has been given before in [5, 25], and we are presenting it here for completeness.</p> <p><b>24</b> That this indicator is that the density distribution of particles is singular. Since for <math>q = 0</math> the original probability distribution is spherically symmetric, it is natural to assume that the associated distribution is spherically symmetric. The radius was computed in [18]. This has also been checked numerically in [6, 7]. A generalisation to all Calde-Yau cone conformal field theory singularities for this result can be found in [25], where it was seen that this procedure matches various other supergravity calculations in the absence of symmetry.</p> <p><b>25</b> For <math>q &gt; 0</math> but very small, we expect that the distributions does not change much, except in the region <math>x^2 \approx 0</math>. Thus we expect that the distribution gives a five dimensional manifold of <math>\mathbb{R}^5</math>, possibly with a boundary, and that the region near <math>x^2 = 0</math> has no density of eigenvalues present. After all, the effective potential repels all particles from the locus <math>x^2 = 0</math>. This</p> | <p><b>18</b> Equations of the saddle point are the following</p> $C - \beta^2 + q \log(a_1^2 + a_2^2) + \int D^2 \phi(x) \log  x  - \beta^2 = 0 \quad (11)$ <p><b>19</b> <math>C</math> is a Lagrange multiplier enforcing the constraint (9). These equations are valid on the support of <math>\rho</math>. We can prove that these equations lead to a singular support for <math>\rho</math>. We do this by contradiction.</p> <p><b>20</b> Assume that a smooth locus for the support of <math>\rho</math> is allowed. In that case, we can differentiate the equation 11 six times, with the Laplacian in 6 dimensions to the cubed power. We find then that</p> $(\nabla^2)^3 q \log(a_1^2 + a_2^2) + \int D^2 \phi(x) \nabla^2 \log  x  - \beta^2 = 0 \quad (12)$ <p><b>21</b> The first two terms obviously are cancelled. Notice that</p> $\frac{\nabla^2(a_1^2 + a_2^2)}{\log(a_1^2 + a_2^2)} \approx \log(x^2) + \log(x^2) \quad (13)$ <p><b>22</b> That away from <math>x^2 = 0</math> we have that</p> $\nabla^2 \log(a_1^2 + a_2^2) = \nabla^2 (\log(x^2) + \log(x^2)) = 0 \quad (14)$ <p><b>23</b> The function is harmonic: a sum of a holomorphic plus an antiholomorphic piece. Also, it is easy to show that <math>(\nabla^2)^3 \log(x^2 - \delta) \propto \delta^2(x^2 - \delta)</math>. With these, we would find that <math>\beta = 0</math>. This argument contradicts that <math>\beta \neq 0</math>, so the assumption that <math>\rho</math> is smooth does not work. This argument has been given before in [5, 25], and we are presenting it here for completeness.</p> <p><b>24</b> That this indicator is that the density distribution of particles is singular. Since for <math>q = 0</math> the original probability distribution is spherically symmetric, it is natural to assume that the associated distribution is spherically symmetric. The radius was computed in [18]. This has also been checked numerically in [6, 7]. A generalisation to all Calde-Yau cone conformal field theory singularities for this result can be found in [25], where it was seen that this procedure matches various other supergravity calculations in the absence of symmetry.</p> <p><b>25</b> For <math>q &gt; 0</math> but very small, we expect that the distributions does not change much, except in the region <math>x^2 \approx 0</math>. Thus we expect that the distribution gives a five dimensional manifold of <math>\mathbb{R}^5</math>, possibly with a boundary, and that the region near <math>x^2 = 0</math> has no density of eigenvalues present. After all, the effective potential repels all particles from the locus <math>x^2 = 0</math>. This</p> | 2ab06e940efc25f7b2e832ab8176de15 |
| 5 | <p><b>DEVOTION TO SERVICE</b></p> <p>For a list of names of people who have been awarded the MGM MIRAGE Corporate Citizenship Award, please visit the website of MGM MIRAGE Corporate Citizenship Award.</p> <p><b>GIVING BACK TO THE COMMUNITIES IN WHICH MGM MIRAGE OPERATES ITS BUSINESSES AND WHERE OUR EMPLOYEES LIVE, WORK, AND CARE FOR THEIR FAMILIES IS A SERIOUS AND DEDICATED COMMITMENT.</b></p> <p><b>OUR COMMITMENT</b></p> <p>MGM MIRAGE is committed to giving back to the communities in which we operate our businesses and where our employees live, work, and care for their families. We are a serious and dedicated commitment.</p> <p><b>OUR COMMITMENT</b></p> <p>MGM MIRAGE is committed to giving back to the communities in which we operate our businesses and where our employees live, work, and care for their families. We are a serious and dedicated commitment.</p> <p><b>OUR COMMITMENT</b></p> <p>MGM MIRAGE is committed to giving back to the communities in which we operate our businesses and where our employees live, work, and care for their families. We are a serious and dedicated commitment.</p>   | <p><b>DEVOTION TO SERVICE</b></p> <p>For a list of names of people who have been awarded the MGM MIRAGE Corporate Citizenship Award, please visit the website of MGM MIRAGE Corporate Citizenship Award.</p> <p><b>GIVING BACK TO THE COMMUNITIES IN WHICH MGM MIRAGE OPERATES ITS BUSINESSES AND WHERE OUR EMPLOYEES LIVE, WORK, AND CARE FOR THEIR FAMILIES IS A SERIOUS AND DEDICATED COMMITMENT.</b></p> <p><b>OUR COMMITMENT</b></p> <p>MGM MIRAGE is committed to giving back to the communities in which we operate our businesses and where our employees live, work, and care for their families. We are a serious and dedicated commitment.</p> <p><b>OUR COMMITMENT</b></p> <p>MGM MIRAGE is committed to giving back to the communities in which we operate our businesses and where our employees live, work, and care for their families. We are a serious and dedicated commitment.</p> <p><b>OUR COMMITMENT</b></p> <p>MGM MIRAGE is committed to giving back to the communities in which we operate our businesses and where our employees live, work, and care for their families. We are a serious and dedicated commitment.</p>  | <p><b>DEVOTION TO SERVICE</b></p> <p>For a list of names of people who have been awarded the MGM MIRAGE Corporate Citizenship Award, please visit the website of MGM MIRAGE Corporate Citizenship Award.</p> <p><b>GIVING BACK TO THE COMMUNITIES IN WHICH MGM MIRAGE OPERATES ITS BUSINESSES AND WHERE OUR EMPLOYEES LIVE, WORK, AND CARE FOR THEIR FAMILIES IS A SERIOUS AND DEDICATED COMMITMENT.</b></p> <p><b>OUR COMMITMENT</b></p> <p>MGM MIRAGE is committed to giving back to the communities in which we operate our businesses and where our employees live, work, and care for their families. We are a serious and dedicated commitment.</p> <p><b>OUR COMMITMENT</b></p> <p>MGM MIRAGE is committed to giving back to the communities in which we operate our businesses and where our employees live, work, and care for their families. We are a serious and dedicated commitment.</p> <p><b>OUR COMMITMENT</b></p> <p>MGM MIRAGE is committed to giving back to the communities in which we operate our businesses and where our employees live, work, and care for their families. We are a serious and dedicated commitment.</p>  | e8d1b976a112cd116d58939049920df  |

Continued on next page

| # | Ground-Truth | SmolDocling | Qwen2.5-VL | ID                               |
|---|--------------|-------------|------------|----------------------------------|
| 6 |              |             |            | eae40317d83ba98d3e70427e6d7baff6 |

Table 2. Visualizations of layout output from SmolDocling and QwenVL-2.5 compared to the DocLayNet ground truth. Examples are chosen to be representative of different layout styles and features. The prediction results however do not represent a generalizable measure of the model’s performance on inputs with similar features. (1) Multi-column pages are handled by SmolDocling and Qwen2.5-VL, with some recall errors in the latter. (2) A manual page with terminal output shows poor bounding box recall on SmolDocling and label confusion in Qwen2.5-VL. (3) Lists with nesting are handled well in SmolDocling but confuse Qwen2.5-VL. (4) Both SmolDocling and Qwen2.5-VL reconstruct equations well, however with different annotation conventions on including or excluding the equation index. (5) On a portrait page with colorful elements and gradient background, SmolDocling creates less accurate bounding-boxes and Qwen2.5-VL suffers low recall. (6) On a report page with tables and diagrams, SmolDocling output exhibits some repetition loop, fabricating non-existent text cells (bottom left), while Qwen2.5-VL is confused between tables and pictures.

## References