

**Group: Sunflower blooming stage classification system**

Group Member:

1. LOH MIN YI 23305383
2. HNEAH GUEY LING 24101308
3. CHAN KHAI SHEN 24101339
4. LIU SIXUAN 23101466
5. CHEN ZIYI 23100909
6. JIANG RUNJIANG 23101516
7. ZHENG CHENGJIN 23101012
8. MOHAMMED JUBARAH 24101399
9. GOH DENG VEE 23101176

Folder: This folder consists of 4 subfolders which are Bud, Partially Bloom, Fully Bloom, and Wilt.

## **Subfolder: Bud**

337 images in the subfolder

### **Goh Deng Vee' s role: Collect 35 images**

For the “Bud” subfolder, I collected pictures from different sources, such as Google website searches, friend’ s contributions, online image libraries of plants and social media platforms. The major problem I encountered was inconsistency of image quality from using different sources and tools. Some photos were blurred and had a lack of consistency of lighting, which made it difficult to notice the features of the bud properly. Also, some of the images had elements of background distractions which included leaves, soil, and other plants thus distracting them from the flower bud itself.

To solve these problems, I carefully filtered out images that were low-quality or blurry. I used Canva to cut out unnecessary elements and scaled all images to be 224 x 224 pixels for continuity. The practice eliminated unnecessary background elements and preserved the picture consistency necessary for the dataset.

#### **Example 1 Before Cropping:**

This picture was retrieved from a Google website through the internet. I selected it because it clearly shows a flower bud with petals tightly enclosed, typical of the early stage of blooming.



#### **After Cropping:**

I used Canva to reduce the size of the image to 224×224 pixels, with the emphasis on the bud and cropping out distracting details that do not relate to the subject.



Example 2 Before Cropping:

This image was provided by a friend, who took a photo of a bud. This picture depicts the usual closed petal configuration of the flower at the early blooming stage.



After Cropping:

Once I sized it down to 224x 224 pixels using Canva I made sure to keep the bud being the focal point and eliminated the background distractions.



Therefore, upon the proper selection and editing of images from various sources, I've managed to build a Bud Subfolder that precisely reflects the early blooming stage of flowers. The images clearly show tightly closed petals, with minimal background distractions and the focus on the buds itself. T

his continued approach makes the dataset clean, organised and ready for analysis.

### **LIU SIXUAN' s role: Collect 114 images**

#### **1. Dataset Sources**

For the “Sunflower Bud” subfolder, I collected images from various everyday platforms to ensure diversity and realism.

##### **1) Xiaohongshu:**

A popular lifestyle-sharing app in China where users often post gardening and plant care content.

You can search for keywords like “Sunflower Planting” or “Sunflower bud”, and you will find many related posts. In the comments section under these posts, other users often share their own experiences and upload additional photos.

##### **2) Weibo :**

One of China' s largest social media platforms. You can search for keywords like “Sunflower Planting” or “Sunflower bud”, and related user posts will appear. Many of these posts have shared photos in the comments from other hobbyists who tried growing sunflowers themselves.

##### **3) Facebook Sunflower Growing Groups:**

I joined sunflower planting communities and collected photos shared by group members.

##### **4) Friends' Personal Gardens:**

Some of my friends who grow sunflowers shared original photos of their plants.

## 5) Taobao Seller Review Sections:

On Taobao, some sellers specialize in selling sunflower buds, either fresh or dried for decoration or culinary purposes. In the product review sections, many buyers upload real photos of the buds they purchased, which became a valuable source of authentic images.

## 2. Challenges Encountered

### 1) Inconsistent Image Quality

Some photos were blurry or out of focus, and the lighting was uneven, making it hard to clearly identify the sunflower buds.

### 2) Messy Backgrounds

Many images contained unrelated objects such as other plants, gardening tools, or cluttered environments, distracting from the main subject.

### 3) Different Image Sizes and Aspect Ratios

Images came in various sizes and shapes, making it difficult to directly use them for model training.

## 3. Solutions

### Step 1: Manual Selection

I carefully checked each image and deleted those that were blurry, low-resolution, or didn't clearly show the sunflower bud.

### Step 2: Cropping Using Phone's Built-in Editor

I used my phone's built-in photo editor to crop out unnecessary backgrounds and ensure the sunflower bud was the main subject.

### Step 3: Batch Resizing Using WPS Office Premium

#### 1): Organizing Images

I created a dedicated folder on my computer and placed all cropped images into it for easy batch processing.

#### 2): Open WPS and Access Image Editor

I opened any image with WPS Office and clicked on “Edit Picture ” in the top menu.

3): Enter Batch Editing

I selected “Batch Edit ” from the toolbar to open the batch processing window.

4): Resize Images to 224×224 Pixels

I imported all images, selected “Resize ”, and set the size to  $224 \times 224$  pixels, meeting the standard input size for machine learning.

5): Save and Overwrite

I selected “Overwrite Original Files” to keep the dataset clean and avoid multiple versions of the same image.

6): All images were standardized to 224×224 pixels with clean backgrounds, making them ready for training and analysis.

#### 4. Team Roles

Myself: Responsible for image collection, manual filtering, cropping, and resizing using WPS Office.

Friends: Provided original sunflower bud photos from their own gardens.

Community Members: Shared planting experiences and sunflower bud images in online communities.

##### Example 1 - Before Cropping

This image was sourced from Weibo, where I found it under a post about sunflower planting by searching keywords like “sunflower bud”. The photo captured a clear sunflower bud in its early growth stage; however, the background was cluttered with other plants and gardening tools, which distracted from the main subject.



### After Cropping

I first used my phone's built-in editor to remove distracting elements and focus on the bud itself. Then, I used WPS Office Premium Version to batch resize the image to  $224 \times 224$  pixels, ensuring the bud stayed at the center of the frame, with background distractions removed.



### Example 2 - Before Cropping

This picture was taken by one of my friends in their personal garden. Although the bud was clearly visible, the image included a wide-angle view that captured excessive empty space and unrelated plants, which diluted the visual focus.



After Cropping

Using my phone' s editor, I tightly cropped the image around the bud to highlight the main subject. Then, I resized it to  $224 \times 224$  pixels using WPS Office' s batch editing tool to keep the dataset consistent.



**CHEN ZIYI' s role: Collect 185 images**

In the process of constructing the "Flower Bud Recognition" dataset, we mainly collected image data from mainstream social platforms, including user real photos under the plant identification topic on Xiaohongshu, pictures po



sted in the comment sections of gardening videos on Douyin, and the super topic content of "What flowers to shoot today" on Weibo. Although these platforms offer abundant image resources, they also bring unique quality challenges. Firstly, the technical issues caused by the platform's characteristics are particularly prominent: the screenshots of short videos have dynamic blur, excessive beautification distorts the structure of the flower center (especially the beauty filters on Xiaohongshu), and the blocky pseudo-images produced by the platform's compression (the secondary compression problem caused by wechat transmission). Secondly, users' shooting habits have introduced complex background interference, including users' hands, price tags in screenshots from e-commerce platforms, and the shadow of the photographer that appears when shooting against the light.

In this project, I adopted a simple and efficient workflow to handle the collected flower bud image data. First, use the "Preview" application that comes with the Apple computer for the initial screening and editing. This step mainly includes two aspects: One is to delete the pictures that obviously do not meet the requirements, such as overly blurry or completely out-of-focus junk pictures; Second, perform basic cropping on the retained pictures to remove irrelevant background distractions and ensure that each picture focuses on the main body of the flower bud. After this round of screening, all qualified pictures will be uniformly saved to a newly created folder. Next comes the batch processing stage. I choose to use WPS Office to complete this task. The specific operation method is as follows: Open the WPS image editor, import all the previously organized images, and then use the "Batch Edit Image Size" function within it. Here, I uniformly adjusted all the pictures to the standard size of 224×224 pixels and selected the option to overwrite the original pictures to save storage space. The entire batch processing process is highly efficient. It usually only takes 1 to 2 minutes to process 100 pictures.

#### Example 1 Before Cropping:

This picture is from Xiaohongshu, a well-known lifestyle sharing platform in China. I deliberately chose this photo because it vividly captures the typical form of the flower bud about to bloom. The petals are closely wrapped, clearly showing the growth characteristics of the flower bud in its early stage.



#### After Cropping:

In this project, I used WPS Office to professionally standardize the collected flower bud pictures. First of all, I checked each picture one by one through the picture editor of WPS, carefully cropped to ensure that the main body of the flower bud was at the center of the picture, and at the same time removed irrelevant background interfering elements. Then, I used the batch editing function of WPS to uniformly adjust all the pictures to the standard size of 224×224 pixels.



#### Example 2 Before Cropping:

This image was sourced from the comment section of a Douyin (TikTok) video, where a user shared their casual observation of a flower bud. The photograph authentically captures the characteristic tightly closed petal formation typical of the initial blooming phase, showcasing how ordinary social media users document plant development in real-world settings.



#### After Cropping:

I standardized all images to  $224 \times 224$  pixels using WPS Office's batch editing feature. During processing, I carefully ensured each bud remained perfectly centered as the focal point while systematically removing any distracting background elements through precise cropping. The WPS resizing tool automatically applied optimal interpolation to preserve critical bud details during dimension adjustment.



### **Subfolder: Partially Bloom**

200 images in the subfolder

### **Hneah Guey Ling' s role: Collect 100 images**

For this assignment, the dataset was primarily collected from online image databases such as Google Images, using specific search keywords like "partially bloomed sunflower." Additional images were gathered from social media platforms like Facebook and Reddit, as well as from agriculture-related web sites where sunflower images are frequently shared. This approach was taken to ensure diversity in the dataset. Images from online databases tend to be more professional, often taken by photographers, while social media platforms provide more natural, real-life photos taken in everyday settings.

Example of image taken from the online database:



Example of image taken from the social media platform:



One of the major hurdles encountered was ensuring the consistency and accuracy of image quality, particularly because sunflowers in the partial bloom stage can vary greatly in appearance. Most people tend to share images of fully bloomed sunflowers, making it difficult to find suitable images that fall exactly within the partially bloomed stage. To address this, each image was manually inspected to filter out blurry or low-quality photos, as well as those showing sunflowers that were either still in the bud stage or already in full bloom. The selection criteria focused on images where the petals were beginning to unfold, showing some green and folded areas, but also displaying enough yellow to clearly indicate partial blooming as shown in the figure below. This careful selection process helped maintain a consistent and relevant dataset for the assignment.

Example of partially-bloomed sunflower which meets the selection criteria:



All images were resized to 224x224 pixels using a Python script to ensure uniformity across the dataset. This preprocessing step was automated to streamline the process and prepare the images for use in the later stages of the project, particularly for training a neural network model for classification tasks. Standardizing the image size helps maintain consistency and compatibility with common input dimensions required by convolutional neural networks.

#### **Loh Min Yi' s role: Collect 100 images**

The data source of these collected images are mainly from social media such as Instagram, RedNote, TikTok, Youtube and Pinterest. The keywords to find the image are: “Sunflower Growing Time Lapse” and “partially bloomed sunflower”. Screenshots are taken from the video when the sunflower is in the stage of partially blooming.

Example of raw image taken from Pinterest :



One of the hurdles are images that contain partially blooming sunflowers are rare as it is not a common image that can be found online. Therefore, to solve this problem, screenshots had been taken from video or reels that contain partially bloomed sunflowers. The image then will crop to focus on the region of interest (ROI). All the images then had been resized to 244 x 244 with python script to ensure consistency across all images.

Example raw Image taken from Youtube:



Example of cropped image:





Subfolder: Fully bloom



277 images in the subfolder

### **Goh Deng Vee' s role: Collect 35 images**

For the “Fully Bloom” subfolder I gathered the images from many sources, including personal photos and going to botanical gardens, contributions from gardening communities, and online collections of floral images. The primary objective was to capture flowers when they are at their peak blooming stage, the flower center clearly visible and with petals fully opened. One of the pictures was taken at my neighbour' s house. I saw a sunflower with all the buds open. I made this shot to emphasize the shape of the flower in its perfect shape with the petals fully open and the center well defined.

One of the difficulties experienced was the variety of flower types because different flowers can look quite distinct when they are in full bloom. That made it difficult to maintain visual consistency for the collection. There were also photos with problems with light, with petals too dark or too bright which reduced visibility of detail. To address these issues, I selected images with clear open petals and distinct centers. I used Canva to fine-tune brightness and contrast for better clarity. Furthermore, I cropped the backgrounds for any irrelevant background and ensured that all images were resized to 224×224 pixels to maintain the dataset consistent. Following these steps, I managed to organize a set of images that properly reflect the bloomed stage.

#### **Example 1 Before Cropping:**

I took the photo below in the house of my neighbour when I saw a sunflower which was completely blooming. The petals of the flower had opened out completely, and the center was visible clearly. This perfectly blooming stage interested me to also include it in my collection.



After Cropping:

To make the image in line with the rest of the collection, I cropped the image using Canva to a 224x224 pixels size thereby isolating the sunflower in question while reducing the amount of distraction.



Example 2 Before Cropping:

This picture was picked up from Quora using a google search. I've picked this photo because it clearly represents a sunflower in full bloom, petals wide open and a flower centre prominent.



After Cropping:

After resizing to 224x224 pixels with Canva, the image now holds the flower at its focus, removing the background.



Therefore, I ensured that the Fully Bloom Subfolder contains high quality images that consistently show flowers at their most vibrant and fully bloomed stage. Each image is organized with consistency in visuals, therefore aiding in correct analysis and representation.

**Mohammed Jubarah role: Collect 176 images**

Fully Bloom Subfolder: Image Collection

Dataset Description: I collected a total of 176 images of sunflowers in full bloom for this subfolder. The primary goal was to gather images showcasing sunflowers at their peak, with fully opened petals and clearly visible centers.

Image Sources: from a variety of online resources to ensure diversity:

- Google Images
- Facebook
- Instagram
- Kaggle
- Roboflow

The intention was to create a dataset with varying lighting, angles, and backgrounds to enhance the robustness of the final model. The dataset includes images with single sunflowers, multiple sunflowers, close-up shots (including some with bees), and images with "noise" such as other flowers or people in the background.

Hurdles and Solutions:

1. Resolution Loss: A significant hurdle was the loss of image resolution when resizing the images to the required 224x224 pixel dimensions. To address this, I attempted to download the highest resolution images available from the source.
2. Time-Consuming Manual Editing: The process of manually cropping and resizing each image was very time-consuming.
3. Inconsistent Aspect Ratios: Many of the original images were rectangular, while the assignment required square images.

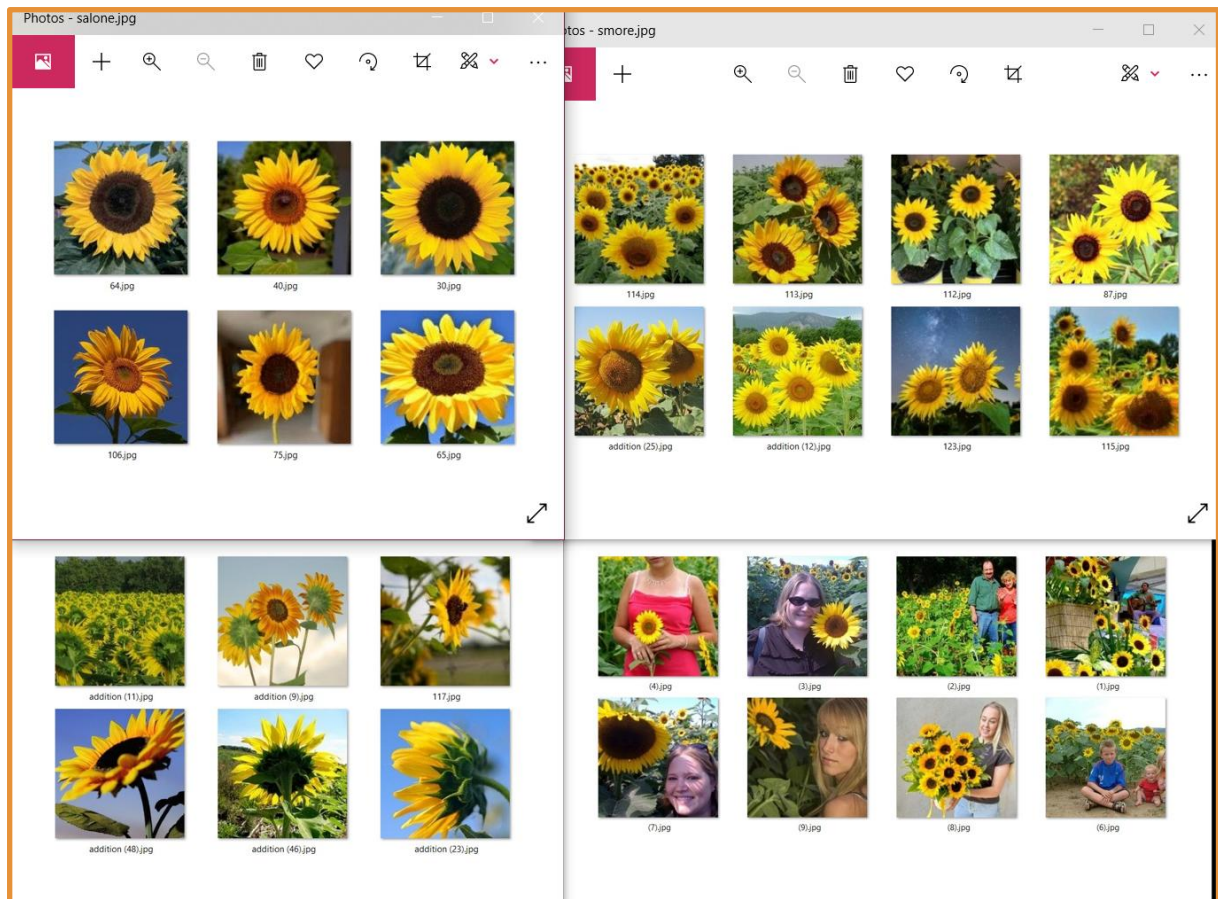
#### Processing Steps:

1. Cropping: All images were initially cropped using the Windows Editor to create a square aspect ratio, focusing on the fully bloomed sunflower and removing extraneous background elements.
2. Resizing: The cropped images were then resized to 224x224 pixels. This was done using a combination of tools:
  - Windows Paint
  - Online image resizing tools (e.g., [iloveimg.com](https://iloveimg.com), [imageresizer.com](https://imageresizer.com))

#### Conclusion:

Through careful collection and processing, a dataset of 176 images of fully bloomed sunflowers was created. While challenges such as resolution loss and time-consuming manual editing were encountered, these were addressed through a combination of careful source selection, cropping, and resizing techniques. The resulting dataset provides a diverse and representative sample of fully bloomed sunflowers suitable for training a multimodal retrieval model.

#### Sample of DS:



Zheng chengjin' s role: Collect 66 images

The sunflower blooming images I collected are mainly sourced from platforms such as Xiaohongshu, Taobao reviews, Google Images, and Google Maps. These channels provide strong support for the diversity and authenticity of the images, helping to comprehensively showcase the natural beauty of sunflowers in full bloom across different regions.

### "Sunflower Blooming" Label (Example from Xiaohongshu)



#### 1. Number of Images

On the Xiaohongshu platform, I collected about 15 high-quality images under the "sunflower blooming" label. The content is rich and covers various angles and scenes.

#### 2. Difficulty in Finding

It is relatively easy to find "sunflower blooming" images on Xiaohongshu. There are many user-shared posts, and searching keywords like "sunflower" or "sunflower field" yields a large number of relevant images. However, it still takes some time and patience to filter out high-resolution, unobstructed, and well-composed photos.

#### 3. Quality and Representativeness of Images

Most of the collected images have good clarity and color performance, accurately reflecting the natural state of sunflowers in bloom. These images show the typical characteristics of blooming sunflowers: complete flower heads, bright colors, and natural backgrounds. They are highly representative samples that meet the needs of learning and classification tasks.



## Quality image in term of sharpness in human eyes and real life photo (10%)

The images collected under the "sunflower blooming" label are generally sharp and clear to the human eye. The details of the sunflower petals and centers are well-defined, and there is minimal blurriness or distortion. Additionally, these images are real-life photographs rather than illustrations or artificial renderings, which ensures that they accurately represent the true appearance of sunflowers in natural settings. This level of sharpness and realism makes the dataset suitable for tasks that require precise recognition and analysis of real-world sunflower characteristics.



images -  
2025-05-17T154237.283



images -  
2025-05-17T154237.342



images -  
2025-05-17T154237.373



images -  
2025-05-17T154237.888



images -  
2025-05-17T154238.102



images -  
2025-05-17T154238.244



images -  
2025-05-17T154238.438



images -  
2025-05-17T154238.614



images -  
2025-05-17T154238.654

My Effort in Collecting and Curating Images



To collect and curate images under the "sunflower blooming" label, I invested considerable effort in several key areas:

#### Searching and Filtering:

I used targeted keywords on the Xiaohongshu platform to search for relevant images. This involved browsing through numerous posts to identify those containing clear and high-quality sunflower photos.

#### Selection Criteria:

I carefully selected images that were sharp, well-composed, and free from significant obstructions or editing filters. Only real-life photographs that accurately represented blooming sunflowers were chosen.

#### Quality Assessment:

Each image was evaluated for its clarity, color accuracy, and representativeness of the learning objective. I excluded images that were blurry, poorly lit, or did not clearly show the sunflower in bloom.

#### Organization:

The final set of images was organized and labeled systematically to ensure they could be easily used for educational or classification purposes.

Overall, the process required attention to detail, patience, and a consistent standard for image quality to ensure that the final dataset was both useful and representative.

## Subfolder: Wilt

235 images in the subfolder

### Chan Khai Shen' s role: Collect 100 images

Among the 100 images, 51 of them are acquired by taking photos of 3 real wilt sunflowers, whereas 49 of them are acquired by collecting images of wilt sunflowers from Google search and Pinterest. For the 51 images acquired from taking photos, the 3 real sunflowers are bought from the florist' s shop when they are blooming, then placed aside to wait for the sunflowers to wilt naturally. The photos are taken from different angles and at different stages of wilting. The 3 sunflowers are shown in the photo below.

Photo of the 3 wilt sunflowers



On the other hand, for the 49 images collected from Google search and Pinterest, they are photos of wilt sunflowers snapped by other people. The search keyword is “wilt sunflower”.

One of the hurdles is that all of the photos taken and images collected are in rectangular shape, but the requirement of the assignment is collecting images in square shape. So, to solve this hurdle, all 100 images are cropped one by one using Microsoft Photos, so that the ratio of length to width becomes 1:1. One example of the image before and after cropping is shown below.

Example of image before and after cropping (rectangle to square)



Another hurdle is that some of the images collected have more than one sunflower. This is considered noise. So, to handle this hurdle, those images are cropped one by one using Microsoft Photos, so that only one sunflower is left in the image. One example of the image before and after cropping is shown below.

Example of image before and after cropping (many sunflowers to one sunflower)



After all 100 images are collected, the images are resized to 224\*224 pixels using Microsoft Photos. All 100 images are in high resolution. The number of pixels of all 100 images before resizing is more than 224\*224.

### **Goh Deng Vee' s role: Collect 35 images**

In the “Wilt” folder, I collected a total of 35 images. My images came from various places including picture repositories of plants, finding photos from gardening forums and choosing from sorted plant image collections. The goal of collecting these images was to capture the late stage of flowering when the petals have faded, drooped, dried, often showing signs of decay and discoloration.

One of the main issues was to catch the proper wilt stage because some of the flowers were only partially wilted while others were totally dried out. Also, in some cases, several wilted flowers were presented, and it added noise and hindered the isolation of one subject. To resolve these, I carefully selected the images to ensure that they accurately depicted the wilted stage, focusing on flowers that clearly showed a darkened floral center and drooping petals. To highlight the areas that are wilted I used Canva to enhance contrast and brightness of an image. To have a dataset that is uniform all images were resized to exact 224x224 pixels.

#### **Example 1 Before Cropping:**

This picture was obtained from an online forum regarding gardening, where users posted photos of their garden sunflowers. I chose it because it demonstrates a clearly wilted sunflower with drooping petals and a bent stem which is characteristic of the final stage of the sunflower' s life cycle.



After Cropping:

To ensure consistency, I cropped the image obtained from Canva to get 224×224 pixels with the focus on the wilted flower and eliminating the background elements.



Example 2 Before Cropping:

This picture was obtained from an online website, Pinterest. I chose it not only because it perfectly illustrates a wilted sunflower and drooping petals and a bent stem, but because the image quality is good enough to make out its indicators of wilting. The photograph effectively captures the natural decline of the flower and thus makes a good representative of the wilt stage.



After Cropping:

Canva was used to shrink the picture to 224x224 pixels so that the wilted flower was the visual emphasis and minimized any distraction of the background.



By sourcing images from online sources and processing them carefully to ensure quality, I made sure that the Wilt Subfolder has uniform and quality images. These photos are a true reflection of the final stage of the life of sunflowers where the emphasis is made on the typical features of wilting such as darkened flower heads and drooping petals.

### **JIANG RUNJIANG' s role: Collect 100 images**

Among these 100 pictures of withered sunflowers, 30 were obtained by photographing a single truly withered sunflower, while the other 70 were collected from pictures of withered sunflowers found on the social media platform Xiaohongshu and the Facebook social group. For these 30 pictures obtained through photography, the single true sunflower was purchased from a flower shop when it was in full bloom, and then placed aside to allow the sunflower to naturally wither. These 30 photos depict a single true sunflower, which was purchased from a flower shop when in full bloom and placed aside to wait for its natural withering. These photos record the state of the sunflowers from different angles and at different stages of withering, as follows.

The photograph of a withered sunflower



Another way is to search for pictures of withered sunflowers online. The 70 pictures collected from Red Cardboard and Facebook are photos of withered sunflowers taken by others and pictures obtained from the sunflower sharing group chats. These photos were obtained through the search "withered sunflowers".

One of the difficulties is that we cannot use cartoon pictures of sunflowers to complete it; instead, we must use real, withered sunflowers.

As shown in the following figure:





The more difficult part is that when obtaining the photos, there will be different sizes, pixels, resolutions and clarity levels. Regarding the issue of size, I think it's similar to cropping to form a square, which doesn't seem particularly out of place. Regarding the issue of clarity, the question is whether the human eye can distinguish whether the photo is clear or not and whether it should be selected.

As shown in the following figure:



Then, during the unified adjustment process, I utilized the WPS image batch processing feature to resize all the photos to a 224 x 224 ratio.