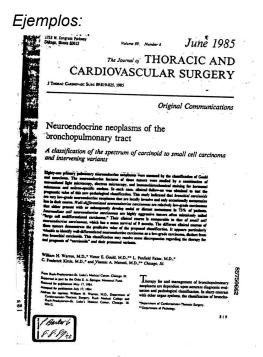
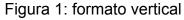


Guía para Ejercicio 5 Entregable 2

Utilizando herramientas vistas en el curso (ej. análisis en frecuencia, transformaciones geométricas, segmentación por umbral, detección de bordes, líneas y regiones, morfología, etiquetado de regiones, entre otros) se buscará identificar en cada imagen:

1. **Orientación general del documento en la imagen**. El documento está en formato vertical o está apaisado.





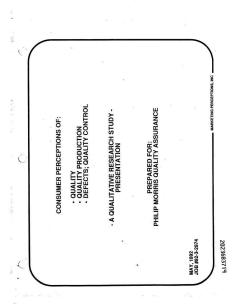


Figura 2: formato apaisado



 Ángulo de giro. Ángulo a girar el documento en sentido horario para que las líneas y/o texto queden horizontales.
 Ejemplo:

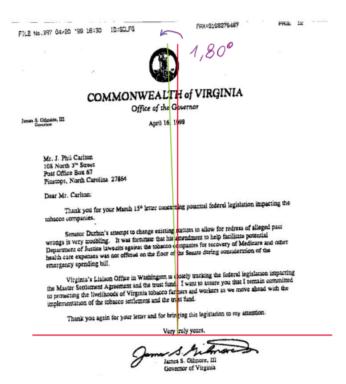


Figura 3: documento original girado (se alinea con giro 1.80° horario)

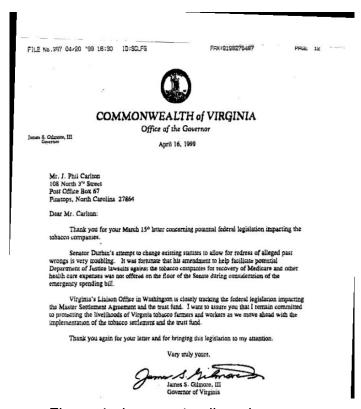


Figura 4: documento alineado



3. ¿El documento es un formulario o tabla? Se entenderá por formulario un documento que cuente con algunas líneas verticales y horizontales (formularios, tablas, etc).

Ejemplos documentos tipo formulario:

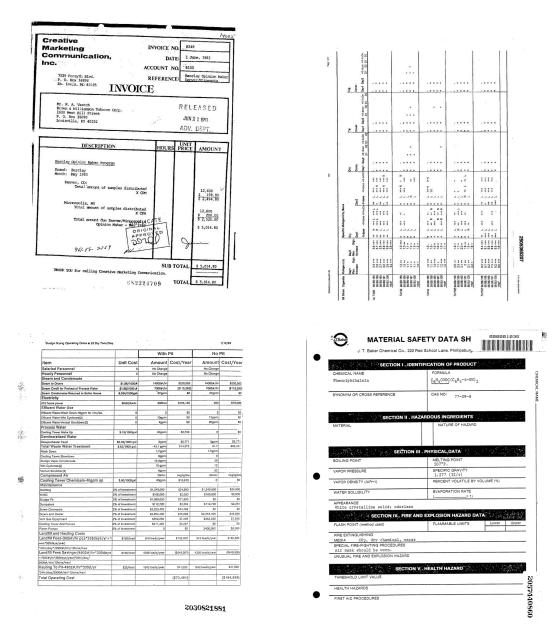


Figura 5: documentos tipo tablas/formularios



4. ¿El documento tiene columnas de texto? El documento tiene agrupaciones de texto en columnas por varias filas.

Ejemplos de documentos con columnas de texto:

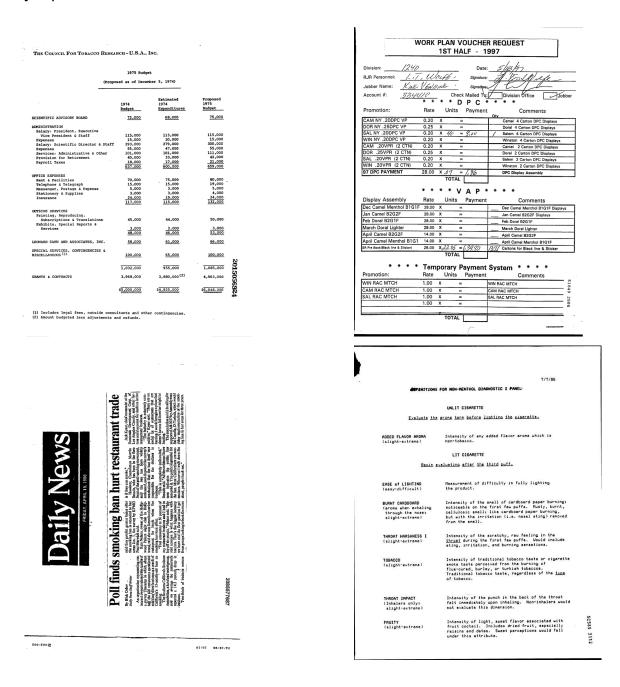


Figura 6: documentos con columnas de texto



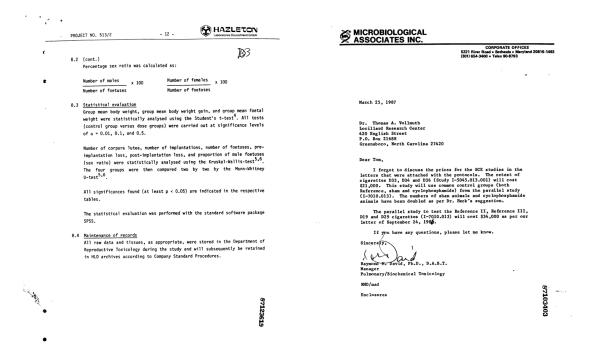


Figura 7: documentos sin columnas de texto

5. **Límites horizontales del documento**. Indicar la columna de inicio y fin del núcleo principal del documento. En caso que (4) sea afirmativo, indicar también las posiciones horizontales que separan las columnas de texto.

Ejemplos de posiciones horizontales. Límites del núcleo principal (rojo) y límites de columnas (celeste), en caso de tener texto en columnas:



| Control Action | Cont

Figura 8: límites (sin texto en columnas)

Figura 9: límites (con texto en columnas)



6. **Alto de línea**. Distancia vertical, en píxeles, entre la línea de base de una fila y la de la siguiente.

Ejemplo:

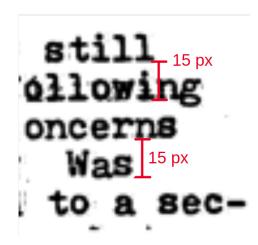


Figura 10: alto de línea

7. **Ancho de letra**. En los casos que el documento tenga un <u>tipo de letra</u> <u>"monospaced"</u> se podrá determinar el ancho de letra como la distancia horizontal en píxeles entre el comienzo de una letra y de la siguiente.

```
ine, Peter Bent Brigham Hospital,
ent in Medicine, Peter Bent Brigh
ete, Laboratory of Biochemical Ge
```

Figura 11: letra monospaced.

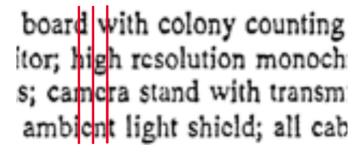


Figura 12: letra proporcional (no monospaced).



8. **Detección de figura**. ¿El documento es sólo texto y/o líneas o contiene alguna figura? Se entenderá por figura un conjunto conexo de píxeles mayor al 1% del área de la imagen y que no está compuesto por texto/líneas y que tampoco es una región quasi-uniforme (ej. un rectángulo quasi-uniforme no se considera una figura)

Ejemplos:







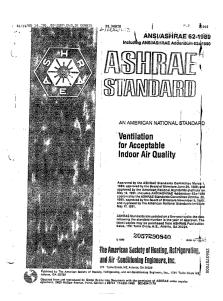


Figura 13: documentos con figuras.