

Guía de escritura de la memoria de un TFG

Pablo Mesejo

Universidad de Granada

Departamento de Ciencias de la Computación e Inteligencia Artificial



UNIVERSIDAD
DE GRANADA



Consideración Previa

- No hay una única manera de elaborar una memoria/informe
 - Ciertas asignaturas o temáticas pueden requerir distintos matices o estructuras particulares

Idea clave

- Un informe o memoria debe cumplir un objetivo fundamental:
 - Explicar
 - qué se ha hecho,
 - cómo se ha hecho lo que se ha hecho,
 - por qué se ha hecho así,
 - qué resultados se han obtenido, y
 - qué conclusiones podemos obtener a partir de los mismos
- Por tanto, **un informe debe ser claro, conciso y bien estructurado.**

Idea clave

- El tribunal, con seguridad, estará muy ocupado y no tendrá tiempo de leer con calma toda la memoria.
- Hay secciones prioritarias (y que deben estar perfectas):
 - Resumen
 - Introducción (Descripción del Problema, Motivación, Objetivos)
 - Conclusiones
- Esas secciones permitirán a cualquier lector, de un vistazo, saber qué se ha hecho en el trabajo, por qué, cómo se ha hecho, y qué resultados principales se han obtenido.

Estructura

Nota: Lo importante es que los contenidos se presenten de modo claro y bien organizado
→ No es necesario seguir siempre al pie de la letra la estructura que presento!

- Portada
 - incluyendo título y autor
- Índice
 - que permita comprender la estructura general de los contenidos y localizarlos con facilidad
 - Por tanto, la memoria debe tener números de página!!

Estructura

1. Introducción

- En donde se describe el problema a resolver (¿qué queremos hacer?), la motivación (¿por qué es relevante hacerlo?), y los objetivos (¿qué objetivos concretos vamos a abordar de cara a resolver el problema?).

2. Fundamentos teóricos

- En donde se presentan los conceptos necesarios para comprender el trabajo.

3. Estado del Arte

- En donde se presenta qué se ha hecho en el campo con anterioridad, y cuáles son los mejores métodos en la actualidad.

Estructura

4. Métodos

- Descripción detallada de los métodos empleados y/o propuestos
- Como una subsección o una sección autónoma podrían incluirse “Detalles técnicos y de Implementación” y/o un apartado de “Planificación”.

5. Experimentos

- Se presentan los datos empleados, el protocolo de validación experimental, las métricas empleadas, los experimentos realizados, los resultados obtenidos, y la discusión de los mismos.
- Dependiendo del TFG, los datos empleados podrían incluirse en otras secciones: una autónoma para los propios datos o en una sección que podría llamarse “Materiales y Métodos”.

Estructura

6. Conclusiones

- Sección que presenta, de modo breve y a modo de resumen, las principales conclusiones del trabajo realizado.
- También suele incluir los trabajos futuros. Es decir, cuáles son las líneas más prometedoras para continuar con este trabajo, así como posibles propuestas de mejora.

7. Bibliografía

- Contiene todas las referencias manejadas por el estudiante a la hora de realizar el trabajo.
- Del texto en el documento se puede ir a consultar la referencia concreta.

8. Anexos

- En caso de que sean necesarios. Sirven para incluir ciertos detalles que se consideren importantes, pero que no se quieran incluir en el cuerpo del documento (para no abrumar al lector, por ejemplo, con numerosas tablas de resultados).

Estructura (ejemplo 1)

Índice general

1. Introducción	15
1.1. Descripción del problema	15
1.1.1. Definición y dificultades del problema	18
1.2. Motivación	20
1.3. Contexto	21
1.4. Objetivos	22
2. Estado del arte	23
2.1. Localización automática de landmarks cefalométricos en fotografías	23
2.2. Localización automática de landmarks cefalométricos de interés odontológico en escáneres y rayos X	25
2.3. Localización automática de landmarks faciales no cefalométricos en fotografías	27
3. Fundamentos teóricos y metodología	37
3.1. Aprendizaje automático y visión por computador	37
3.1.1. Aprendizaje automático	37
3.1.2. Aprendizaje supervisado y no supervisado	37
3.1.3. Problemas de regresión y clasificación	38
3.1.4. Visión por Computador	39
3.2. Aprendizaje por máxima pendiente	39
3.2.1. Función de coste o pérdida	40
3.2.2. Gradiente descendiente	41
3.3. Deep Learning	42
3.3.1. Perceptron Multicapa o Redes Neuronales <i>feed-forward</i>	42
3.3.2. Back-propagation	43
3.3.3. Optimizadores y AMSGrad	44
3.4. Redes Neuronales Convolucionales Profundas	45
3.4.1. Convolución	46
3.4.2. Pooling	47
3.4.3. Batch normalization	48

3.4.4. Relación de las capas convolucionales y la profundidad de la red	48
3.4.5. Ajuste fino o <i>fine-tuning</i>	48
3.5. <i>Data augmentation</i> y manipulación de imágenes 2D y modelos 3D	49
3.5.1. <i>Data augmentation</i>	49
4. Datos y métricas	51
4.1. Conjuntos de datos del problema	51
4.1.1. Descripción	51
4.1.2. Limpieza de errores	53
4.1.3. Ejemplos 3D	55
4.2. Métricas del problema	56
5. Implementación	59
5.1. Diseño del software	59
5.2. Entorno de ejecución	62
6. Experimentos	63
6.1. Separación de conjuntos y validación de los modelos	63
6.2. Decisiones experimentales	63
6.2.1. Framework de ejecución	64
6.2.2. Elección del modelo base	66
6.2.3. Descripción de HyperFace-Resnet101	67
6.2.4. Optimizador elegido	71
6.2.5. Generación de las proyecciones 3D	72
6.2.6. Utilización del dataset original y el 3D	73
6.2.7. Determinación de hiperparámetros	74
6.2.8. Fine-tuning y preentrenamiento sobre AFLW	76
6.2.9. Online Data-Augmentation	76
6.3. Proceso de entrenamiento	77
6.3.1. Evaluación de los resultados de 5-fold CV	79
6.4. Dificultades encontradas durante la experimentación	79
6.5. Resultados	81
6.5.1. Combinación dataset original y modelos 3D	81
6.5.2. Impacto del optimizador elegido	83
6.5.3. Comparación de los resultados con el conjunto de test	84
6.5.4. Análisis del mejor modelo elegido	86
6.5.5. Problema de regresión	96
6.5.6. Problema de clasificación	98
7. Conclusiones y trabajos futuros	99
Bibliografía	103

Estructura (ejemplo 2)

Índice general

Índice de figuras	III
Índice de cuadros	V
1. Introducción	1
1.1. Descripción del problema	1
1.2. Motivación	4
1.3. Objetivos	5
2. Fundamentos Teóricos	7
2.1. Aprendizaje automático y aprendizaje profundo	7
2.2. Redes Convolucionales Profundas	8
2.2.1. Convolutional layer	9
2.2.2. Pooling layer	10
2.2.3. Fully-connected layer	10
2.3. Técnicas de aumento de los datos	11
2.4. Métodos clásicos	11
3. Estado del Arte	15
3.1. Clasificación de imágenes con <i>deep learning</i>	18
3.2. Enfoques de IA para la determinación del sexo a partir de imágenes óseas	18
4. Planificación e Implementación	21
4.1. Planificación	21
4.2. Implementación y entorno de ejecución	22
5. Datos	23
5.1. Descripción de los datos	23
5.2. Preprocesado de los datos	24
6. Métodos Propuestos	27
6.1. <i>Deep learning</i>	27

6.2. Técnicas Clásicas de Aprendizaje Automático y Visión por Computador	30
7. Experimentos	33
7.1. Aumento de datos	33
7.2. Protocolo de validación experimental	34
7.3. Métricas	35
7.4. Experimentación con <i>deep learning</i>	37
7.5. Experimentación con Técnicas Clásicas de Aprendizaje Automático y Visión por Computador	38
7.6. Comparativa Global con Experto Humano y Estado del Arte	39
8. Conclusiones	41
9. Bibliografía	43

Estructura (ejemplo 3)

Índice general

1. Introducción	15
1.1. Descripción del problema	15
1.2. Motivación	17
1.3. Objetivos	18
2. Estado del arte	21
3. Fundamentos Teóricos y Metodología	25
3.1. Resonancia Magnética Funcional	25
3.1.1. La señal BOLD	25
3.1.2. Balloon model	28
3.2. Computación Evolutiva	32
3.2.1. Definición	33
3.2.2. Algoritmos evolutivos	33
3.2.3. Componentes de los algoritmos evolutivos	35
3.3. Algoritmos meméticos	39
3.3.1. Estructura general de los algoritmos meméticos	40
3.3.2. Meta-operador de mutación	40
4. Planificación, Desarrollo y Entorno de Simulación	43
4.1. Planificación	43
4.2. Diseño e Implementación	45
4.2.1. Traducción a GNU Octave	45
4.2.2. Herramientas de ejecución de experimentos	47
4.2.3. Herramienta de transformación y análisis de resultados	47
4.3. Entorno de simulación	48
5. Métodos Propuestos	53
5.1. Expectation-Maximization/Gauss-Newton	53
5.1.1. Recursive EM/GN	54
5.1.2. Randomized EM/GN	54
5.2. Metaheurísticas consideradas para resolver el problema de estimación de parámetros del Balloon model	56
5.2.1. Differential Evolution	56
5.2.2. Real Coded Genetic Algorithms	57
5.2.3. Particle Swarm Optimization	59

5.2.4. Covariance Matrix Adaptation Evolution Strategy	61
5.2.5. Algoritmos Meméticos	62
5.3. Función Objetivo	63
6. Experimentos	65
6.1. Configuración de los métodos	65
6.2. Conjuntos de datos considerados	66
6.3. Descripción de los tests estadísticos	68
6.4. Resultados	70
6.4.1. Experimentos con Datos Sintéticos	70
6.4.2. Experimentos con Datos Reales	74
6.4.3. Experimentos Complementarios	83
7. Conclusiones	89

Estructura (ejemplo 4)

Índice general

1. Introducción	1
1.1. Descripción del problema	1
1.2. Motivación	4
1.3. Objetivos	5
1.4. Estructura de la memoria	5
2. Fundamentos teóricos	7
2.1. Machine learning y deep learning	7
2.2. Redes neuronales	8
2.2.1. Capa totalmente conectada	10
2.2.2. Capa Convolutiva	11
2.2.3. Capa de Pooling	12
2.2.4. Dropout	13
2.3. Neural Architecture Search	13
2.3.1. Espacio de búsqueda	14
2.3.2. Estrategia de búsqueda	17
2.3.3. Estrategia de estimación del rendimiento	22
3. Estado del arte	25
3.1. Resumen estado del arte	25
3.2. Evolución de NAS	25
3.2.1. Aproximaciones evolutivas	25
3.2.2. Aprendizaje por refuerzo	28
3.2.3. Enfoque jerárquico	29
3.2.4. Optimización bayesiana	30
3.2.5. NAS en la actualidad	30
4. Materiales y métodos	31
4.1. Datos del problema	31
4.2. Métricas	35
4.3. Función de pérdida	38
4.4. Red diseñada por experto	38
4.5. Aprendizaje por refuerzo: ENAS	40

ii

ÍNDICE GENERAL

4.6. Optimización bayesiana: Auto-Keras	42
4.7. Algoritmo Evolutivo: Auto CNN	43
5. Planificación e Implementación	47
5.1. Planificación	47
5.2. Implementación	48
5.3. Lenguaje y entorno	50
6. Experimentación	51
6.1. Red diseñada por experto	53
6.2. ENAS	53
6.3. Auto-Keras	56
6.4. Auto CNN	58
6.5. Discusión	59
7. Conclusiones y Trabajos Futuros	63
Bibliografía	65

Cuestiones a evitar

- Introducir resultados (tablas, figuras) y no comentarlos.
 - Resultados sin discutir → es como no presentar resultados.
- Introducir texto de modo innecesario.
 - Evitar “andarse por las ramas” y meter párrafos “por rellenar”.
 - Intentar responder con claridad las hipótesis científicas manejadas o las preguntas planteadas.
- Afirmaciones sin justificación empírica o teórica.
 - **Cuando afirmamos algo es porque los resultados de nuestros experimentos o la literatura científica existente nos permiten afirmarlo.**

Cuestiones a evitar

- Sobre introducir texto de modo innecesario.

Es interesante discutir, ligeramente puesto que no es parte del ámbito de estudio de este proyecto, algunas de las ideas tomadas para la obtención de los materiales. Para la muestra AM la muestra suele estar limitada al uso de fotografías, la razón detrás de esto es por la facilidad de su obtención. Es muy común poder disponer de una variedad de fotografías relativas a una persona desaparecida, pero en cambio es muy improbable contar con un modelo 3D de la cabeza para su uso.

VS

La muestra AM suele estar limitada al uso de fotografías. Es muy común disponer de una gran variedad de fotografías relativas a una persona desaparecida, mientras que es improbable contar con un modelo 3D de su cabeza.

Recomendaciones

- No deja de ser un TFG de Ingeniería, por lo que **se debe incluir alguna sección dedicada a planificación** (temporal y económica) **e implementación**
 - En planificación económica, hay que presupuestar el trabajo por horas, pero también el material empleado.
 - Integrar, dentro de lo posible, diagramas de clases, diagramas de casos de uso, diagramas de secuencia, diagrama de arquitectura o paquetes del proyecto, y/o diagramas E/R.
 - Compartid vuestro código en GitHub. ¡Importante!
- Será valorado positivamente por aquellas personas más interesadas en desarrollo e ingeniería del software.

Recomendaciones

- **Fondo y forma son ambos importantes**
 - Emplead LaTeX y las plantillas proporcionadas por la ETSIIT
 - Logo de la UGR actualizado, nombre de tutores, titulación
- **Numerad** tablas y figuras, **e incluid pies de tabla/figura explicativas**. Referenciad también en el texto las tablas y figuras.

Recomendaciones

- Evitar memorias esquemáticas.
 - Se deben **discutir y analizar los resultados obtenidos** (tanto a nivel cualitativo como cuantitativo).
 - No vale decir “Podemos ver el resultado durante la ejecución del programa” o “Podemos ver el resultado en la tabla” (sin indicar nada más)
 - Es recomendable **incluir información que ponga en valor el trabajo realizado**:
 - ventajas e inconvenientes de los métodos empleados,
 - problemas encontrados y cosas que se han aprendido durante la realización del TFG,
 - conocimientos adquiridos en asignaturas que han sido empleados,
 - número de experimentos lanzados y duración de los mismos,
 - etc.

Recomendaciones

- A la hora de presentar el estado del arte, vale la pena presentar un histograma (elaborado en *Scopus*, por ejemplo) de publicaciones en el campo
 - Da una medida de la “importancia” de ese campo de investigación.
 - Es necesario **indicar el buscador, la *query* empleada y la fecha de consulta.**

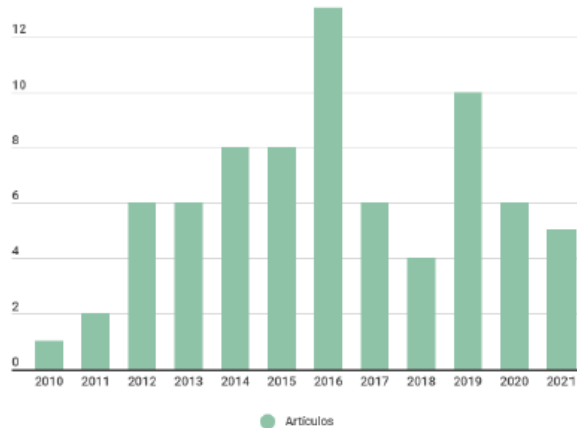


Figura 20: Histograma que representa los artículos por año encontrados a través de la herramienta *Scopus*.



Esta figura es claramente mejorable:

- No se sabe qué consulta (con qué palabras clave) se ha empleado
- No se sabe cuándo se realizó esa consulta.

Recomendaciones

- Las figuras y tablas deben verse correctamente
 - **Una figura o tabla que no se ve, o no se entiende, no cumple su función.** Es como si no estuviera...
 - **Limitar el número de decimales** en las tablas (2 o 3 suelen ser suficiente) → aumenta la legibilidad de los resultados
 - Se recomienda encarecidamente explicar conceptos, resultados y métodos de modo visual → **Una imagen vale más que mil palabras!**
- Integrar las fórmulas y ecuaciones en el propio documento (y numerarlas)
 - No pegarlas/incrustarlas como una imagen en medio del texto!

Recomendaciones

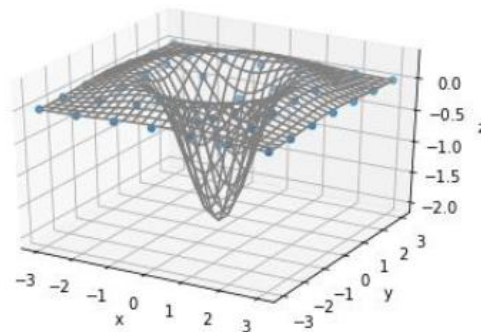
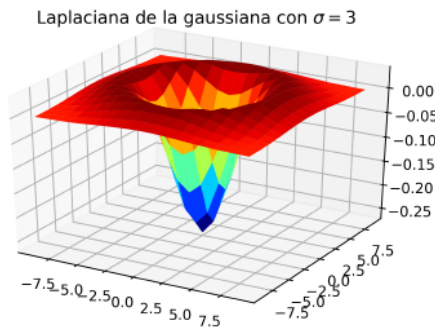
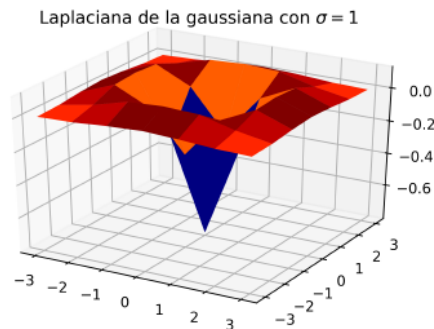



Figura 1.5: En el caso de $\sigma = 1$, el tamaño de la máscara es 7, y al ser tan pocos puntos no se aprecia del todo bien la forma de la laplaciana. En el caso de la derecha, con una máscara de tamaño 19, ya se aprecia bien la forma de sombrero mejicano invertido.

Esta imagen tiene pie de imagen (*caption*), está numerada y es muy clara

Esta imagen no tiene pie de imagen (*caption*), no está numerada y la resolución de la misma es mejorable

Nota: si las figuras no son creación vuestra, debéis indicar la fuente original en la *caption*.

Recomendaciones

- En términos globales, se parte de lo general y se va a lo particular:

Mayor grado de concreción

 1. Presentamos el problema a resolver y por qué es interesante y difícil resolverlo
 2. Presentamos los conceptos fundamentales para comprender el trabajo que el lector está a punto de comenzar a leer
 3. Presentamos el estado del arte, es decir, qué han hecho otros investigadores antes que nosotros, y cuáles son las mejores técnicas a día de hoy en base a la evidencia empírica
 4. Presentamos los materiales y métodos que vamos a usar en nuestro trabajo: datasets, métodos empleados o diseñados por nosotros
 5. Presentamos los experimentos realizados, los resultados obtenidos, y la discusión de los mismos.
 6. Presentamos las conclusiones y trabajos futuros.
- Desde este punto de vista, no tiene sentido, por ejemplo, presentar un concepto general (p.ej., la definición de cross-validation) en la mitad de los experimentos. Debería ir al comienzo de los experimentos, como parte de la presentación del protocolo de validación experimental o, en todo caso, cuando se hable de materiales y métodos (si, en esa sección hablamos también, por ejemplo, de métricas de evaluación del trabajo).

Recomendaciones

- La bibliografía es importante.
 - Muestra hasta qué punto se han cuidado todos los detalles del trabajo y su acabado final.
 - Se debe incluir toda la información que permita asignar autoría y localizar la información:
 - Autores, nombre de revista, año de publicación, volumen, número y números de página.
 - Se use el estilo/formato que se use, siempre debe ser el mismo para todas las referencias → consistencia



Lathuilière, S., Mesejo, P., Alameda-Pineda, X., & Horaud, R. (2019). A comprehensive analysis of deep regression. *IEEE transactions on pattern analysis and machine intelligence*, 42(9), 2065-2081.



Lathuiliere, S., Mesejo, P., Alameda-Pineda, X., & Horaud, R. (2019). A comprehensive analysis of deep regression. *IEEE*, 42(9).

El nombre de un autor no es correcto (falta una tilde)

El nombre de la revista no está completo

No se incluyen números de página

Recomendaciones

- Guiad/ayudad al lector lo máximo posible, facilitándole el trabajo.
- La pregunta que nos debemos hacer es:
Si vosotros no hubieseis hecho esta memoria, al leerla, ¿comprenderíais el trabajo desarrollado?

Guía de escritura de la memoria de un TFG

Pablo Mesejo

Universidad de Granada

Departamento de Ciencias de la Computación e Inteligencia Artificial



UNIVERSIDAD
DE GRANADA

